

Topology Control in Wireless Ad hoc Networks

by

Ahmed Mohamed Ali Omer Hassan

*Thesis presented in fulfilment of the requirements for the
degree of Master of Science in Mathematics at Stellenbosch
University*



Department of Mathematical Sciences,
Faculty of Science,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Supervisor: Prof. J.W. Sanders

December 2013

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the owner of the copyright thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature:
Ahmed Hassan

Date: 2013/12/27

Copyright © 2013 Stellenbosch University
All rights reserved.

Abstract

Topology Control in Wireless Ad hoc Networks

Ahmed Hassan

*Department of Mathematical Sciences,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.*

Thesis:

December 2013

Wireless *ad hoc* networks are increasingly used in today's life in various areas ranging from environmental monitoring to the military. For technical reasons, they are severely limited in terms of battery power, communication capacity and computation capability. Research has been carried out to deal with these limitations using different approaches. A theoretical treatment of the subject is topology control whose basic task is to design network topologies with special properties that make them energy-efficient and interference-optimal.

We study, implement and compare the XTC and CBTC algorithms in terms of interference reduction, length stretch factor and maximum degree. These two algorithms have two features that are absent in almost all competitive topology control algorithms which are practicality and maintaining connectivity. Both algorithms show good performance in terms of interference reduction and maintaining a good length stretch factor. Regarding CBTC, we prove that it is a power spanner.

We show through extensive simulation that the degree distribution of wireless *ad hoc* networks modelled by the log-normal model is binomial if the average degree is not high. We find that there is no fixed threshold for the average degree at which the distribution is distorted and no longer binomial.

We show through simulation that the node density which ensures the absence of isolated nodes is a tight lower bound for the node density which ensures connectivity. The implication of this result is that connectivity is ensured with high probability if the minimum node degree is equal to 1.

Finally we show through simulation that the log-normal model is not a realistic representation of wireless *ad hoc* networks if the environmental parameter is at least 6. This result is important because there are no available

ABSTRACT

iv

measurements to determine the range of the environmental parameter for typical frequencies used in wireless *ad hoc* networks.

Uittreksel

Topology Control in Wireless Ad hoc Networks

Ahmed Hassan

*Departement Wiskundige Wetenskappe,
Universiteit van Stellenbosch,
Privaatsak X1, Matieland 7602, Suid Afrika.*

Tesis:

Desember 2013

Koordlose *ad hoc* netwerke word toenemend gebruik in vandag se lewe op verskillende gebiede wat wissel van die omgewing monitor tot militêregebruik. Vir tegniese redes is hulle ernstig beperk in terme van battery krag, kommunikasie kapasiteit en berekeningsvermoë. Navorsing vanuit verkillende benaderings word gedoen om met hierdie beperkings te deal. 'n Teoretiese benadering tot onderwerp is topologie beheer. Die basiese taak is om netwerktopologieë met spesiale eienskappe wat hulle energie-doeltreffend en interferensie-optimaal maak te ontwerp.

Ons bestudeer, implementeer en vergelyk die XTC en CBTC algoritmes in terme van interferensie vermindering, lengte rek faktor en maksimum graad. Beide hierdie algoritmes het twee eienskappe wat afwesig is in byna al die mededingende topologie beheer algoritmes: hulle is prakties en handhaf verbindings. Beide algoritmes toon goeie prestasie in terme van interferensie verminder en die handhawing van 'n goeie lengte rek faktor. Ten opsigte van CBTC bewys ons dat dit 'n "power spanner" is.

Ons wys deur middel van uitgebreide simulatie dat die graad verdeling van die koordlose *ad hoc* netwerke wat deur die log-normale model gemodelleer kan word binomiaal is as die gemiddelde graad nie hoog is nie. Ons vind dat daar geen vaste drempel is vir die gemiddelde graad waarby die verdeling vervorm en nie meer binomiaal is nie.

Ons wys deur simulatie dat die node digtheid wat die afwesigheid van geïsoleerde nodusse verseker 'n streng ondergrens vir die node digtheid wat konektiviteit verseker is. Die implikasie van hierdie resultaat is dat 'n konneksie verseker word as die minimum node graad gelyk is aan 1.

Ten slotte wys ons deur simulاسie dat die log-normale model nie 'n realistiese voorstelling van koordlose *ad hoc* netwerke is wanneer die “environmental parameter” groter is as 6 nie. Hierdie resultaat is belangrik, want daar is geen beskikbare metings om die grense van hierdie parameter vir 'n tipiese frekwensie gebruik in koordlose *ad hoc* netwerke te bepaal nie.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Prof. J.W. Sanders, for his supervision, encouragement and support. I am very privileged to work with him.

Dedications

To

My parents:

Mohamed & Sakina

Contents

Declaration	ii
Abstract	iii
Uittreksel	v
Acknowledgements	vii
Dedications	viii
Contents	ix
List of Figures	xi
List of Tables	xiv
Nomenclature	xv
1 Introduction	1
1.1 Wireless <i>Ad hoc</i> Networks	1
1.2 Problem Statement	1
1.3 Problem Significance	2
1.4 Related Work	2
1.5 Thesis Structure	6
2 Theoretical Foundation	7
2.1 Glossary of Graph Theory	7
2.2 Modelling Wireless <i>Ad hoc</i> Networks	10
2.3 Random Graph Model	10
2.4 Radio Channel Models	11
3 Topology Control	18
3.1 Definition of Topology Control	18
3.2 Criteria of Good Topology	19
3.3 Taxonomy of Topology Control	21
3.4 Topology Control and Interference	22

CONTENTS

x

4	Non-homogeneous Topology Control	27
4.1	XTC	27
4.2	Cone-Based Topology Control Algorithm	32
4.3	Features of XTC and CBTC	46
5	Simulation	49
5.1	Degree Distribution in Wireless <i>Ad hoc</i> Networks	49
5.2	Connectivity and Minimum Node Degree	62
5.3	The Range of the Environmental Parameter	67
6	Conclusion	71
6.1	Summary	71
6.2	Future work	72
	Bibliography	74

List of Figures

2.1	Links in relative neighbourhood graphs	9
2.2	Yao links	9
2.3	Log-normal graph.	16
	(a) $\zeta = 1$ and $r_0 = 1.0$	16
	(b) $\zeta = 1$ and $r_0 = 1.5$	16
2.4	From (Woo <i>et al.</i> , 2003).	17
	(a) Reception rate of all links.	17
	(b) The radio coverage area.	17
3.1	Taxonomy of topology control techniques.	21
3.2	Different scenarios of interference	23
	(a) Hidden terminals.	23
	(b) Exposed terminals.	23
	(c) Asymmetric terminals.	23
3.3	Link interference	23
3.4	Optimum topology with non constant interference.	25
3.5	Interference.	26
	(a) Two exponential chains.	26
	(b) Relative neighbourhood graph.	26
	(c) Optimum topology.	26
3.6	Locality can conflict with interference reduction	26
4.1	XTC applied to a network.	29
	(a) Max-power graph.	29
	(b) XTC.	29
4.2	Non constant stretch factor	31
	(a) When n is even	31
	(b) When n is odd	31
4.3	Maximum degree of G_{XTC}	31
4.4	Crossing links	32
4.5	A cone of angle θ	33
4.6	Illustration of shrink-back operation.	34
4.7	Asymmetric links	35
4.8	Illustration of Lemma 4.2.	36

4.9	A disconnected topology	37
4.10	Two nodes in the same cone	38
4.11	A path connecting u and v in G_θ	39
4.12	A special triangle	40
4.13	The node v is below $\overline{u_{k-1}u_k}$	40
4.14	The node v is above $\overline{u_{k-1}u_k}$	41
4.15	Neighbours in the same cone at the same distance	42
4.16	CBTC applied to a network with $\theta = \pi/2$	42
	(a) max-power graph.	42
	(b) CBTC.	42
	(c) CBTC (asymmetric links are removed).	42
4.17	The spanner property is violated	43
4.18	The angle $\angle vwu$ is acute.	44
4.19	Power stretch factor	45
4.20	Interference is reduced when CBTC is used.	46
4.21	Interference of XTC and CBTC.	47
	(a) With Interference of max-power graph	47
	(b) Without Interference of max-power graph	47
4.22	Length stretch factor.	47
4.23	Maximum degree.	48
5.1	Normal approximation to binomial distribution	56
	(a) Bad approximation	56
	(b) Good approximation	56
5.2	Probability mass function	58
	(a) $n = 100$, $\zeta = 2$ and $r_0 = 1.4$	58
	(b) $n = 100$, $\zeta = 1$ and $r_0 = 1.8$	58
5.3	Connectivity improves as ζ increases.	60
	(a) $\zeta = 1$	60
	(b) $\zeta = 2$	60
	(c) $\zeta = 3$	60
5.4	Relation between ζ and r_0	60
5.5	Presence and absence of border effect (a)	61
	(a) Border effect is present	61
	(b) Border effect is avoided	61
5.6	Presence and absence of border effect (b)	61
	(a) Border effect is present	61
	(b) Border effect is avoided	61
5.7	Effect of r_0 on the degree distribution	61
	(a) $r_0 = 1.0$	61
	(b) $r_0 = 2.0$	61
	(c) $r_0 = 3.0$	61
5.8	Minimum node degree and connectivity	65
	(a) $ A = 100$, $n = 200$, $\zeta = 2$ and $r_0 = 2.0$	65

(b) $ A = 100$, $n = 200$, $\zeta = 3$ and $r_0 = 2.0$	65
5.9 Relation between ρ and r_0	66
5.10 Link probability	68
5.11 Local clustering coefficient	69
5.12 Clustering coefficient	69
5.13 Giant component size	70

List of Tables

2.1	Values of path loss exponent	13
5.1	Link probability	54
5.2	Degree distribution	57
5.3	Minimum node density	66

Nomenclature

Variables

β	The path loss	[dB]
α	The path loss exponent	[]
\mathcal{P}	The predetection threshold	[dB]
P_a	The area mean power	[dB]
σ	The standard deviation of variations in signal strength	[dB]
ζ	The environmental parameter	[dB]
θ	The parameter of the CBTC algorithm	[rad]
ρ	The node density	[m ⁻²]
r	L^2 -norm	[m]
c	The power stretch factor	[]
r_0	The distance at which $P_a = \mathcal{P}$	[m]
A_0	The radio coverage area	[m ²]
λ	The intensity of Poisson distribution	[]

Chapter 1

Introduction

This chapter briefly introduces wireless *ad hoc* networks, specifies our problem in its very general sense, shows its significance, presents the related work and finally shows the general structure of the thesis.

1.1 Wireless *Ad hoc* Networks

Wireless *ad hoc* networks are types of networks in which network units (nodes) communicate through wireless channels; hence the term “wireless”. In addition they do not have a fixed infrastructure and thus they are very easy to deploy for purposes that arise on-the-fly; hence the term “*ad hoc*”. In wireless *ad hoc* networks there are no central nodes which implies that all nodes have the same status. Each node can be a source of information, a destination to which the information is to be sent or act like a router to forward the information towards the intended destination. Each node is equipped with a battery, memory, transceiver, computational unit and sometimes sensor. If the nodes are equipped with sensors we refer to the network as a wireless sensor network (WSN). WSNs are deployed, usually, to monitor a condition of interest in a particular observational site. They are frequently used in today’s life; for example they have been used in environment to monitor the weather and to control the temperature in intelligent buildings, in agriculture to control the irrigation system and in the military to detect opponent intrusion. If the nodes are mobile we refer to the network as a mobile *ad hoc* network (MANET). MANETs have many applications in real-life situations; for instance they have been used in rescue mission to keep the rescue team members in contact.

1.2 Problem Statement

Due to technological reasons wireless *ad hoc* networks have limited resources in terms of power, computational and communication capabilities.

These limitations have triggered research attempting optimization of the algorithms and protocols running on wireless *ad hoc* networks. This objective can be achieved through many techniques such as designing power-aware routing protocols, etc. A theoretical technique for achieving this objective is *topology control*.

The primary task of topology control is to construct network topologies that prolong the network lifetime through minimizing energy consumption and increase the network throughput through confining interference. To explain how topology control achieves these two goals, we need firstly to model the network as a *graph* called a *max-power* graph. It models a situation in which each node uses its maximum transmitting power. As a consequence a good deal of energy is wasted and interference is very high. Topology control protocols are special algorithms that run on a max-power graph and compute a subgraph called a *reduced topology* which satisfies certain properties of interest.

In short, our research question is how to construct topologies for wireless *ad hoc* networks that minimize energy consumption and maximize network throughput.

1.3 Problem Significance

Although it is still mainly of interest of theoreticians, topology control has been proven to achieve its objectives through many simulation studies. Its significance is clear since it is technologically and economically infeasible to manufacture network devices with unlimited power and computational and communication capabilities. An effective way to overcome these limitations is through extensive research aimed at constructing network topologies that optimize resource-usage. Indeed, topology control is one of these research directions.

1.4 Related Work

For a general view of topology control we recommend the extensive survey (Santi, 2005b). It surveys the work done in *homogeneous topology control* as well as *non homogeneous topology control* in the context of stationary wireless *ad hoc* networks as well as mobile wireless *ad hoc* networks. In fact, it is more than just a survey because it even presents a popular model used for mathematically representing wireless *ad hoc* networks, the tools used to study them and ends with open issues.

This thesis contributes to homogeneous and non homogeneous topology control. Therefore we survey and discuss the closely related work done in both areas.

The research question in homogeneous topology control is finding the minimum transmitting range that guarantees connectivity of the network. This minimum transmitting range is referred to as the *critical transmitting range* (CTR). The motivation of this research question is the reported fact (Santi, 2005b) that the network capacity will be compromised if each node uses its maximum transmitting range. In addition, usually nodes are battery-equipped with no online power supply and in some cases even with no way of recharging the batteries. Consequently, it is highly desirable to reduce energy consumption by setting to each node the minimum transmitting range that is just large enough to ensure connectivity of the network.

The pioneer work in this area in 2-dimensions is done in (Gupta and Kumar, 1998). It is shown therein that if nodes are placed uniformly in a disc of unit area in \mathbb{R}^2 and each node sets its transmitting range to $\sqrt{[\log n + c(n)] / \pi n}$ the network is connected with high probability (w.h.p) if $c(n) \rightarrow \infty$ as $n \rightarrow \infty$ where n is the number of nodes. This result is proved when the *border effect* is ignored and when it is considered.

In (Wan and Yi, 2004) a stronger notion of connectivity has been studied: *k-connectivity*. The paper obtains the CTR that ensures the asymptotic *k*-connectivity w.h.p in a unit square and a unit circle with and without the border effect with uniform spatial distribution.

Though very interesting, the practical value of these results is unclear (Santi and Blough, 2003; Stuedi *et al.*, 2005) because they hold only asymptotically, i.e., when $n \rightarrow \infty$, and consequently they are applicable only for extremely dense networks. In reality, it is very unlikely to set up a very dense network because of issues regarding the network capacity and performance.

(Santi and Blough, 2003) circumvent this problem by obtaining an upper bound for the CTR that ensures connectivity w.h.p for *sparse* networks and validating the analytical expression through simulation. The novelty of (Santi and Blough, 2003) is making the length, l , of the side of the deployment area an additional parameter in the problem. Having the CTR as a function of n and l allows the obtained results to apply to dense *as well as* sparse networks. By simulation, the paper studies the case when nodes are mobile. In this case the CTR is interpreted as the minimum transmitting range that ensures connectivity for a certain period of time. Two mobility models are considered: the random way point (RWP) model which represents intentional motion and a Brownian-like model which represents unintentional motion. It has been found that in both cases mobility reduces connectivity. This result is also reported in (Bettstetter, 2004b). Surprisingly, the analysis of connectivity with these two completely different models shows an overall agreement. This implies that the difference between the analysis of connectivity in stationary networks and mobile networks is solely due to mobility *irrespective* of the detailed description of the mobility pattern. This point needs to be formally investigated because it is invaluable for simulation studies. If the detailed description of mobility models does not really matter then a simple Brownian-like model can be used

instead of more sophisticated models. In addition, research in mobility models might become less active.

In (Bettstetter, 2002) a lower bound for the CTR that ensures k -connectivity w.h.p is obtained and verified by simulation. This work has been generalised in (Bettstetter, 2004b) where connectivity is studied when the spatial distribution is arbitrary, uniform, Gaussian and RWP with and without the border effect. Three different perspectives of connectivity has been analysed: *i*) Connectivity from the node-wide perspective. In this case the question is: what is the CTR that ensures the existence of no isolated node? *ii*) Connectivity from the path-wide perspective. Since there is no known analytical expression for the probability that two given nodes are connected by a path in 2-dimension, the problem is studied through simulation. *iii*) Connectivity from the network-wide perspective.

The research in k -connectivity is motivated by the fact that wireless *ad hoc* networks are failure-prone. One way to design a failure-resilient network is to ensure k -connectivity which guarantees that the network will not fail unless at least k nodes fail. At the other extreme, (Ta *et al.*, 2009; Santi and Blough, 2003; Hekmat and Van Mieghem, 2006) have shown that if connectivity is relaxed and instead we require only a large fraction of the network, say 90% of the network, to be connected, huge energy can be saved. In particular, it is proved (Ta *et al.*, 2009) that the ratio of the minimum transmitting range that ensures a large fraction is connected to the CTR tends to zero as the number of nodes tends to infinity implying that as the number of nodes increases the amount of energy saved as a result of relaxing connectivity increases. This result is confirmed in (Hekmat and Van Mieghem, 2006) via simulation. However in (Santi and Blough, 2003) it has been shown through simulation that the ratio between the two transmitting ranges approaches 0.5. This seems to contradict the analytical result obtained in (Ta *et al.*, 2009). Fortunately we can justify this apparent contradiction. In (Santi and Blough, 2003) the transmitting ranges are obtained for sparse networks while the result of (Ta *et al.*, 2009) is applicable only for very dense networks. Moreover, they consider two different network models.

In mobile wireless *ad hoc* networks, relaxing connectivity can be interpreted as ensuring connectivity for a certain period of time instead of ensuring connectivity all the time. The reported results imply that large amounts of energy can be saved if temporary dis-connectivity can be tolerated.

It is not straightforward to decide which one of these two extreme views, ensuring k -connectivity versus relaxing connectivity, is preferred. First of all, they are still theoretical. In addition, it somehow depends on the objective of setting up the network. For example in the military and security, it is imperative to maintain connectivity all the time while in monitoring environmental conditions it is usually sufficient to require 90% connectivity to enable nodes to send their collected data to the intended destination.

All the aforementioned literature assume a simple network model as they

assume that networks are modelled as random geometric graphs. The random geometric graph model ignores variations in the received signal strength which definitely occur in reality. When these variations are considered we have a more realistic model called the log-normal model. In (Bettstetter and Hartmann, 2005) connectivity is analysed using the log-normal model. A tight lower bound for the node density required for ensuring connectivity with a certain probability is obtained using a multidimensional integral and validated through simulation. This result is useful in simulation for planning wireless sensor networks deployed in a region with a known size because we can figure out how many nodes are to be placed in the deployment region to guarantee connectivity with a given probability. In (Hekmat and Van Mieghem, 2006) connectivity is analysed using the same model, yet it is merely a simulation study apart from one analytical expression obtained for the giant component size of wireless *ad hoc* networks as a function of the average degree. A novel aspect of (Bettstetter and Hartmann, 2005) is that ensuring connectivity with a given probability defines a “level” of connectivity.

(Santi, 2005a) provides the first analytical results for the CTR for mobile *ad hoc* networks. The results are derived for the RWP as well as for any mobility model provided that it describes the motion of the nodes within a defined area and the probability of finding a mobile node at the boundaries is zero. Fortunately, these two assumptions are common in mobility models. In the RWP, nodes pause for a certain period of time, p , after reaching their destinations. If $p = 0$, i.e., nodes are mobile all through the simulation time, an approximation for the CTR is found in terms of the CTR for stationary networks with uniform spatial distribution. The approximation shows the existence of an asymptotic gap between the CTR for mobile networks and the CTR for stationary networks. In (Santi and Blough, 2003) it is reported through simulation that as the number of mobile nodes increases, connectivity is decreased.

Non-homogeneous topology control is an active area of research which can be divided into many active sub-areas. One of them is the study of techniques used to construct network topologies in a fully distributed and local manner. Based on the type of information required to compute the reduced topology, non-homogeneous techniques can be classified into three groups: location-based, direction-based and neighbourhood-based.

The location-based techniques require exact information about node positions. An example of this technique is (Li *et al.*, 2003) which computes an approximation of MST locally; hence it is named LMST. The position information can be provided by a global positioning system (GPS). Using GPS, however, consumes node energy remarkably. Moreover the locations of the nodes are more likely to change in wireless *ad hoc* networks and that makes the location-based techniques not robust to mobility. Therefore we have chosen not to include the location-based techniques in this thesis.

The direction-based techniques require each node to determine directions

of its neighbours. The pioneer work in these techniques is the cone-based topology control algorithm (CBTC) introduced in (Wattenhofer *et al.*, 2001) which computes a variant of the Yao graph; see §.2.1. The topology produced by the CBTC is proved to preserve connectivity and to be a power spanner, see §.3.2. The CBTC is improved in (Li *et al.*, 2005) by adding a set of optimizations. These optimizations reduce transmitting ranges without violating connectivity. In these two papers the network is modelled using the random geometric graph model which implies that links are symmetric. In (Xiang-Yang *et al.*, 2005) a similar approach is used when the network is modelled by a *mutual inclusion graph* in which two nodes are connected if they are within the transmitting range of each other. This approach is useful if links are not symmetric. However, considering asymmetric links does not improve network performance; rather it might increase the overhead of routing protocols which are typically designed under the symmetry assumption (Santi, 2005b).

The neighbourhood techniques are the cheapest in terms of the type of required information for computing the reduced topology as they require each node only to be able to identify its neighbours and order them based on a certain criterion. The XTC introduced in (Wattenhofer and Zollinger, 2004) builds a variation of the relative neighbourhood graph, see §.2.1, in a completely distributed and local manner. The topology produced by the XTC is always connected, planar, and has a bounded degree. A similar protocol to the XTC is the k -Neigh protocol (Blough *et al.*, 2003) which builds a topology by letting each node choose at most k nearest neighbours based on their distance. The topology computed by the k -Neigh protocol is connected w.h.p. This protocol is inspired by the theoretical result derived in (Xue and Kumar, 2004) regarding the number of nearest neighbours with which each node has to connect to guarantee asymptotic connectivity with probability one.

1.5 Thesis Structure

This thesis is organized as follows:

The second chapter introduces the foundations on which the other chapters are based. Particularly, it introduces two widely used models for wireless *ad hoc* networks.

The third chapter defines topology control, mentions the criteria of good topologies, classifies the topology control techniques and discusses interference reduction through topology control.

The fourth chapter studies the XTC and CBTC algorithms from a mathematical viewpoint and compares them in terms of interference, length stretch factor and maximum degree.

The fifth chapter studies, by simulation, the degree distribution of wireless *ad hoc* networks, connectivity and minimum degree and the effect of the environmental parameter on the log-normal model.

Chapter 2

Theoretical Foundation

This chapter presents the theoretical foundations required for better comprehension of the theoretical treatment of wireless *ad hoc* networks presented in later chapters. The objective of the chapter is to introduce two models of wireless *ad hoc* networks that are extensively used in the *topology control* literature. The models are ultimately *graphs* and hence the key terms of graph theory used within this thesis are briefly introduced in §.2.1. The suitability of random graph model for representing wireless *ad hoc* networks is concisely discussed in §.2.3. In §.2.4, we present two models for radio propagation. Modelling radio propagation is necessary to model wireless *ad hoc* networks reliably as we shall see that the choice of radio propagation model determines the wireless *ad hoc* network model. Finally the two models of wireless *ad hoc* networks are presented in §.2.4.1 and §.2.4.2.

2.1 Glossary of Graph Theory

In this section we introduce some terms borrowed from *graph theory* and used throughout this thesis. There are variety of books introducing graph theory and studying it from different viewpoints. We have chosen to rely on (Diestel, 2005).

A *graph* is a pair of two sets $G = (V, E)$ where $E \subset V \times V$. The elements of V are called *nodes* and the elements of E are called *links*. The *cardinality* of V is called the graph's *order* and is denoted by $|V|$. Wherever we use n it refers to $|V|$ unless stated otherwise.

Links can have orientation (direction). In this case E is not symmetric. When links have direction we refer to the graph as *directed* graph. Unless stated otherwise, we assume that all graphs we deal with are *undirected*.

To each link we can assign a real number which is called *weight*. In wireless networks, a link's weight can represent the cost of a wireless link which is defined in §.3.2.

The nodes u and v are *neighbours* if $(u, v) \in E$. The set of neighbours of

u , denoted by N_u , is defined as $\{v \in V \mid (u, v) \in E\}$. The cardinality of N_u is called the *degree* of u which we denote by d_u . A node is called *isolated* if $d_u = 0$. If there exists $(u, u) \in E$ we call it a *loop*. The graphs we deal with here have no loops. The average degree of a graph G is defined as

$$\bar{D} = \frac{1}{n} \sum_{v \in V} d_v. \quad (2.1.1)$$

Let $G' = (V', E')$. If $V' \subset V$ and $E' \subset E$ then G' is called a *subgraph* of G . On the other hand, G is a *supergraph* of G' . A subgraph G' is called *spanning* if $V' = V$ and $E' \subset E$.

A *path* between u and v (or from u to v), denoted by $\Pi(u, v)$, is a sequence of nodes u_1, u_2, \dots, u_k with $u_1 = u$, $u_k = v$ and $(u_i, u_{i+1}) \in E$ for $i = 1, 2, \dots, k-1$. It is called a *cycle* if $u_1 = u_k$ and $k > 2$. The length of a path is the number of links that form the path. Two paths connecting u and v are said to be *independent* if they do not have a node in common except u and v .

A graph is *connected* if there is a path between any pair of nodes. A graph is called *k-connected*, where k is a positive integer, if removing any set of less than k nodes will not result in disconnecting the graph. Alternatively, a graph is *k-connected* if any pair of nodes is connected by k independent paths.

In a disconnected graph, if there exists a component whose order is much bigger than all other component's order we refer to it as a *giant component* (Bollobas, 2001).

A *tree* is a graph which is connected and has no cycles. A *spanning tree* is a tree that spans all the nodes, i.e., includes all the nodes. A *minimum spanning tree* (MST) is, then, a tree that connects all nodes and its total link weight is at most the total link weight of any other spanning tree.

A *sparse graph* is a graph with $|E| = O(n)$, i.e., the number of links is linear in n .

A graph is embedded in a d -dimension Euclidean space if there is a map, $F : V \rightarrow \mathbb{R}^d$, that associates each node in V with a point in \mathbb{R}^d . All graphs we consider in this thesis are embedded in a 2-dimensional region.

A *planar graph* is a graph embedded in a plane in such a way that it does not contain crossing links.

One of widely used types of graph in the theoretical study of wireless *ad hoc* networks is the *random geometric graph* (Hekmat, 2006a; Wiki01, 2013) which is constructed by distributing nodes independently at random according to some probability distribution, most often uniform distribution, and connecting two nodes if and only if they are at most a threshold r_0 apart. The term “geometric” is because it is assumed that the graph is embedded in a geometric space, most often \mathbb{R}^2 . The term “random” is rather weird in this context because links are formed in a very deterministic way, however there is a source of randomness which is the distribution of nodes which makes a link event between two given nodes random.

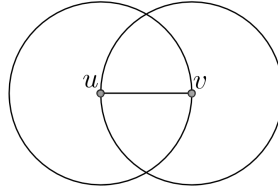


Figure 2.1: Since the intersection of the two circles is empty, the link (u, v) is formed.

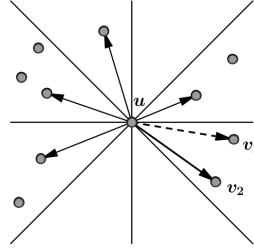


Figure 2.2: Node u establishing its Yao links. The Plane is divided into 8 separate cones. In each cone the node connects to a nearest neighbour if any. Note that v_1 and v_2 is at the same distance from u but u has chosen to form a link with v_2 and ignored v_1 . This process is called *tie breaking*.

Many geometric structures in computational geometry can be constructed *locally*, hence they attract researchers to make use of them in building topologies for wireless *ad hoc* networks. A graph is constructed locally if each node builds its own links without requiring global knowledge about other nodes. Well-known locally constructable geometric structures include the relative neighbourhood graph and the Yao graph.

In a *relative neighbourhood graph* (RNG) (Buchin and Buchin, 2007) a link between two nodes exists if and only if they are relatively closest to each other. More precisely, in RNG a link (u, v) is established if the intersection of two circles centred at u and v with radius equal to the distance between u and v is empty. The intersection of these two circles is called the *relative neighbourhood*. An illustration is shown in Figure 2.1.

A *Yao graph* (Yao, 1982), denoted by $Y\vec{G}_k$, is constructed by letting each node divide the plane into k separate cones around it, where k is a positive integer. In each cone the node is connected to a nearest neighbour with a directed link pointing to the neighbour. If the directions of the links are ignored we get the undirected Yao graph denoted by YG_k ; see Figure 2.2.

2.2 Modelling Wireless *Ad hoc* Networks

At any instant of time, a wireless *ad hoc* network can be modelled as a graph, often called the *communication graph*, $G = (V, E)$ where the elements of V are the network units which live in a 2-dimensional region.¹ A link between two nodes exists if and only if there is a wireless radio channel between them. A wireless radio channel between two nodes is established if and only if the transmitted signal is received above a minimum power threshold (Hekmat, 2006a; Santi, 2005c; Ta *et al.*, 2009; Bettstetter and Hartmann, 2005). We assume that all radio channels are symmetric², i.e., if node v correctly receives a message sent by node u , then node u receives a message sent back by node v using the same transmitting power. We simply mean links are two-way or bidirectional or undirected. It is worth mentioning that we consider two nodes connected based on the requirement mentioned above regardless of whether the two nodes can actually transfer data or not. In other words, we separate network topology from network capacity and thus if at any instant of time two nodes cannot transfer data to each other due to interference, congestion, etc. we still consider them connected as long as our condition of connectivity is met. This is in accordance with the OSI model as it separates the *physical layer* which is responsible for establishing connections between two network nodes regardless of whether the connection is reliable or not from the *data layer* which is responsible for a reliable connection through which data is guaranteed to be transferred between two network units. In addition, separating network topology from network capacity eases the theoretical treatment of connectivity.

The graph which models the case when each node uses its maximum transmitting power is referred to as the *max-power graph*. Sometimes we call it the *original topology* because it represents the network topology in the absence of topology control.

2.3 Random Graph Model

Random graphs are very popular and widely used to model real-life networks. In this section we answer the following question: what is the suitability of random graphs for modelling wireless *ad hoc* networks?

It should be noted that the model of random graphs we are discussing here is the Erdős-Rényi model. There are two equivalent variants of the Erdős-Rényi model for generating random graphs (Bollobas, 2001, Ch. 2). In the first model, we start with a set consisting of n nodes and we choose M links

¹In general, the deployment region is assumed to be a d -dimensional region where $d = 1, 2, 3$.

²This assumption is common in the literature (Ta *et al.*, 2009) and it is important because many distributed algorithms running on the network require acknowledgement of messages.

independently and uniformly from $N = \binom{n}{2}$ possible links. In this case there are $\binom{N}{M}$ equiprobable random graphs that can be constructed in this way with M links. In the second model, each possible link occurs independently with probability p . The expected number of links is then Np . Throughout this thesis the term “random graph” refers to the graph constructed using the second model.

From the definition of the random graph model, it is deduced (Hekmat, 2006a) that it is not a realistic representation of wireless *ad hoc* networks for the following reasons:

1. In wireless *ad hoc* networks, links depend on distance, i.e., the probability of having a direct link between any two nodes decreases as the distance separating them increases. In random graphs, all links equally likely occur irrespective of the distance.
2. In wireless *ad hoc* networks, links are *locally correlated*. Links are said to be locally correlated if the probability that two randomly chosen nodes are connected given that they have a common neighbour is higher than the probability that two randomly chosen nodes are connected given that they do *not* have a common neighbour. In random graphs, links are independent and thus these two events equally likely occur.

The link dependence on distance and link correlation are vitally important features of wireless *ad hoc* networks and since the random graph model fails to capture those features it is not a realistic model for wireless *ad hoc* networks.

2.4 Radio Channel Models

In wireless *ad hoc* networks communications are carried out through radio channels. A transmitted signal suffers, on its way to the receiver, from attenuation and it is correctly received only if it is above a power threshold, called the *predetection threshold*, when received at the receiver.

Modelling signal propagation is of theoretical and practical value. From the theoretical viewpoint deciding whether we establish a link between two nodes in the graph modelling the network depends heavily on the model used to describe signal propagation as we shall see in §.2.4.1 and §.2.4.2. From the practical point of view the network functionality and the quality of service are negatively affected by link impairments which are due to signal attenuation through the radio channel.

The mechanisms that govern signal propagation make it quite difficult to predict the signal strength eventually received at the receiver. These mechanisms can generally be classified into three categories (Andersen *et al.*, 1995): *i)* *reflection* which occurs when the signal hits an object whose dimension is very large when compared with the wavelength of the signal, *ii)* *diffraction*

which takes place when the signal hits an object with very sharp edges and *iii) scattering* which occurs when the signal hits many objects whose dimension is small compared with the signal's wavelength.

Fortunately there are reliable models that predict the strength of the received signal. Two of them will be the subject of §.2.4.1 and §.2.4.2. The third model predicts variations in signal strength over a very short period of time (order of milli-seconds up to a few seconds) or over very short distances (order of a few wavelengths). It is assumed that small scale strength variations are distributed according to either the Rician or Rayleigh distribution depending on whether we considering Line-of-Sight or not. Taking this model into account may lead to more accurate theoretical results, nonetheless it renders the theoretical approach too complicated and thus is ignored.

As previously mentioned, the signal attenuates in the radio channel as it is propagating. The attenuation of the signal is referred to as the *path loss* which can be defined as (Santi, 2005c), (Santi, 2005b), (Bettstetter and Hartmann, 2005)

$$\beta = \frac{P_t}{P(r)}, \quad (2.4.1)$$

where P_t is the transmitting power and $P(r)$ is the received strength which depends on the distance r between the transmitter and the receiver in a given surrounding environment.

Eq.(2.4.1) can be expressed in terms of decibels³ (dB) as follows

$$\beta[\text{ in dB }] = 10 \log_{10} \left(\frac{P_t}{P(r)} \right).$$

Generally the path loss, β , can be decomposed into two components (Bettstetter and Hartmann, 2005; Hekmat, 2006a): a deterministic component which will be presented in §.2.4.1 and a stochastic component. The stochastic component can be further decomposed into two parts: small scale strength variations which we decide to ignore and medium scale strength variations which follow the log-normal distribution and will be presented in §.2.4.2.

2.4.1 Path Loss Model

This model is obtained based on measurements and curve fitting (Hekmat, 2006a; Labrador and Wightman, 2009). The measurements were taken as follows: a transmitter is fixed in a specific location and the received signal strength is measured at many locations with the *same* distance to the transmitter. Although the distance to the transmitter is the same, there are variations in the received signal strength. These variations in the received signal strength are referred to as medium scale strength variations. Then the received

³Decibel is a logarithmic unit used to expressed the ratio between two quantities such as intensity. The ratio P_0/P_1 can be expressed in dB as $10 \log_{10} (P_0/P_1)$.

Environment	α
free space	2
urban area	2.7 – 3.5
indoor (LoS)	1.6 – 1.8
indoor (no LoS)	4 – 6

Table 2.1: Typical values of the path loss exponents. LoS stands for Line-of-Sight. From (Santi, 2005c, p. 16).

signal strength is averaged over many locations with the same distance to the transmitter over many times.⁴ The average strength is referred to as the *area mean power*. The distance to the transmitter is kept changing and the whole experiment is repeated. It is obvious that the path loss model ignores medium scale strength variations.

Measurements have shown that the area mean power attenuates as the power law of the distance separating the transmitter and the receiver. Mathematically, the area mean power can be approximated by (Hekmat, 2006a; Labrador and Wightman, 2009; Santi, 2005c; Bettstetter and Hartmann, 2005)

$$P(r) = P_t P_0 \frac{d_0}{r}^\alpha, \quad (2.4.2)$$

where P_t is the transmitting power, P_0 is the strength received at a reference distance d_0 , r is the distance separating the transmitter from the receiver and α is the *path loss exponent*. The path loss exponent captures the environment in which the network is deployed. High values of α correspond to environments with a lot of obstructions and irregularities. Typical values of α are shown in Table 2.1.

Eq.(2.4.2) can be expressed in terms of dB as follows

$$P(r) [\text{in dB}] = 10 \log_{10} (P_t) + 10 \log_{10} (P_0) - 10\alpha \log_{10} \frac{r}{d_0}. \quad (2.4.3)$$

By combining Eq.(2.4.1) and Eq.(2.4.2), the path loss is given by:

$$\beta \propto r^\alpha. \quad (2.4.4)$$

As we mentioned in §.2.2 a wireless link is established between two nodes if and only if the received signal strength at the receiver is above a threshold which we called the predetection threshold. Let \mathcal{P} denote the predetection threshold. Then according to the path loss model we have

$$\mathcal{P} = P_t + P_0 - 10\alpha \log_{10} \frac{r_0}{d_0}, \quad (2.4.5)$$

⁴Even at a fixed location there are variations in the signal.

where \mathcal{P}, P_t, P_0 are measured in dB and r_0 is the distance at which the area mean power equal to the predetection threshold.

As a consequence of Eq.(2.4.5) all nodes at most r_0 away from the transmitter correctly receive the transmitter's signal and thus according to the path loss model the radio coverage area is a perfect disk centred at the transmitter with radius equal to r_0 .

If we denote by \mathcal{L} the link event, i.e., two randomly chosen nodes establish a link, then

$$P[\mathcal{L}|r] = \begin{cases} 1 & \text{if } r \leq r_0 \\ 0 & \text{if } r > r_0, \end{cases} \quad (2.4.6)$$

where $P[\mathcal{L}|r]$ is the probability that two randomly chosen nodes are directly connected given that they are r distance apart.

From Eq.(2.4.6) it is obvious that a wireless *ad hoc* network modelled by the path loss model is equivalent to a random geometric graph in which the threshold is r_0 and nodes are distributed according to some probability distribution function.

Random geometric graphs are widely used in the literature to model wireless *ad hoc* networks partly for the following reasons: *i*) they are simple, *ii*) links depend on distance, and *iii*) links are locally correlated.

In reality, links in wireless *ad hoc* networks are distance-dependent and locally correlated.

2.4.2 Log-normal Model

As we mentioned in §.2.4.1 according to the path loss model the received strength is obtained by averaging the signal strengths. Thus the path loss model ignores variations in the strength of the received signal. In reality, depending on the complexity of the environment, the variation in the signal strength at locations with the same distance to the transmitter might be prominent and hence ignoring them might render the path loss model unreliable.

The basic assumption of the log-normal model is that the logarithmic value of the received strength is normally distributed around the logarithmic value of the area mean power given by Eq.(2.4.3) with standard deviation σ . Therefore according to the log-normal model the received signal strength can be approximated by (Andersen *et al.*, 1995; Bettstetter and Hartmann, 2005; Hekmat, 2006a; Hekmat and Van Mieghem, 2006; Ta *et al.*, 2009):

$$P(r)[\text{in dB}] = P_t + P_0 - 10\alpha \log_{10} \frac{r}{d_0} + Z, \quad (2.4.7)$$

where P_t and P_0 are as in Eq.(2.4.3) but measured in dB and Z is a random variable following the normal distribution with mean 0 and standard deviation σ . The standard deviation σ represents the fluctuations in the strength of the received signal caused by obstructions and irregularities in the environment.

Measurements have shown that σ can be as large as 12 in very obstructed environments (Hekmat, 2006a).

When $\sigma = 0$ (no shadowing) the log-normal model reduces to the path loss model. Although the log-normal is more realistic and a generalization of the path loss model it is less frequently used in the literature compared with the path loss model, perhaps owing to the simplicity provided by the latter model.

As we know the signal is correctly received if $P(r) \geq \mathcal{P}$ which implies

$$P(r) = P_t + P_0 - 10\alpha \log_{10} \frac{r}{d_0} + Z \geq \mathcal{P} = P_t + P_0 - 10\alpha \log_{10} \frac{r_0}{d_0} ,$$

and that gives

$$Z \geq 10\alpha \log_{10} \frac{r}{r_0} ,$$

therefore we have

$$P[\mathcal{L}|r] = P[P(r) \geq \mathcal{P}] = P\left[Z \geq 10\alpha \log_{10} \frac{r}{r_0}\right] ,$$

where $P[E]$ means the probability of the occurrence of the event E . Since Z is normally distributed with mean 0 and standard deviation σ ,

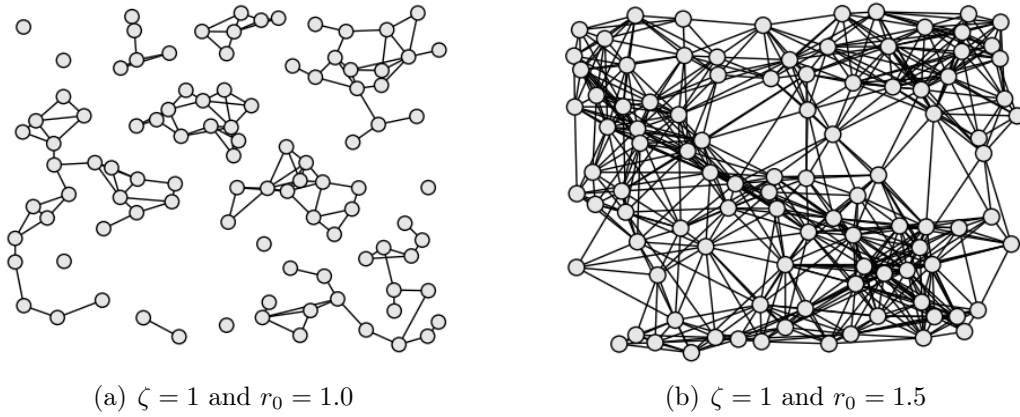
$$\begin{aligned} P[\mathcal{L}|r] &= \frac{1}{\sqrt{2\pi}\sigma} \int_{10\alpha \log_{10}(\frac{r}{r_0})}^{\infty} \exp\left[-\frac{z^2}{2\sigma^2}\right] dz \\ &= \frac{1}{2} \operatorname{erfc} \left[\frac{z}{\sqrt{2}\sigma} \right]_{10\alpha \log_{10}(\frac{r}{r_0})}^{\infty} \\ &= \frac{1}{2} \left[1 - \operatorname{erf} \left[\frac{10}{\sqrt{2}} \frac{\log_{10}(r/r_0)}{\zeta} \right] \right] , \end{aligned} \quad (2.4.8)$$

where $\zeta := \sigma/\alpha$ and $\operatorname{erf}(\cdot)$ denotes the *error function* which is defined as (Spiegel *et al.*, 2000):

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-u^2} du.$$

We call ζ the *environmental parameter* since it reflects the environment. Note that when $\zeta = 0$, Eq.(2.4.8) reduces to Eq.(2.4.6) which confirms that when the variations in signal strength are ignored the log-normal model reduces to the path loss model. It is convenient to abstract the environment in one parameter and thus when we consider the log-normal model we only talk about ζ instead of talking about σ and α individually.

Generally, links established according to Eq.(2.4.8) are not symmetric. To be consistent with our assumption of link-symmetry we can either remove

**Figure 2.3:** Log-normal graph.

directed links or make them undirected by adding the other direction to the graph. The first procedure will produce a sparse graph while the second will produce a dense graph.

We refer to a graph whose links are constructed according to Eq.(2.4.8) when the undirected links are made directed by adding the other direction as a *log-normal graph*. Figure 2.3 shows examples of log-normal graphs where each point represents a network unit and each link represents a wireless link.

In a log-normal graph, because of the probabilistic nature of Eq.(2.4.8) the following two events are possible:

1. Removal of a link whose length is less than r_0 . We refer to such a link as a *short* link.
2. Emergence of a link whose length is larger than r_0 . We refer to such a link as a *long* link.

Consequently the radio coverage area is no longer a perfect disk. In reality, the radio coverage area is not a perfect disk due to the fluctuations in the received signal strength.

It is worth mentioning that an empirical study done in (Woo *et al.*, 2003) shows that *i*) two pairs of nodes with the same separating distance can have a very different link quality.⁵ Therefore the very deterministic view of the link-dependency on distance assumed by the path loss model is, of course, a simplification of reality. *ii*) In addition, some of the relatively close nodes to the transmitter have poor link quality while some of the more further nodes have good link quality. This point provides support for the notion of removal of some short links and emergence of some long ones presented by the log-normal model.

⁵In (Woo *et al.*, 2003) link quality means the rate of packet reception. We can use $P[\mathcal{L}|r]$ as an abstraction of link quality. If $P[\mathcal{L}|r] \approx 1$, the link is of good quality whereas if $P[\mathcal{L}|r] \approx 0$ the link is of poor quality.

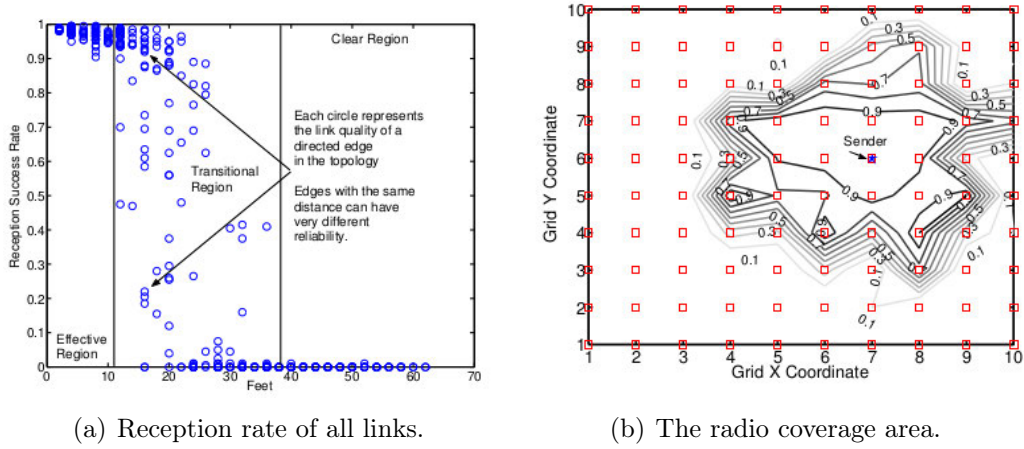
**Figure 2.4:** From (Woo *et al.*, 2003).

Figure 2.4(a) illustrates the two points mentioned above. It shows the scatter plot of link quality represented as points and plotted against distance measured in feet. In the figure the *effective region* is a region where the link quality is very high, i.e., very good connectivity, the *clear region* is a region where the link quality is very low, i.e., very poor connectivity and the *transitional region* is a region in between the two regions. In the transitional region the variation in the received signal strength is very high which explains why the link quality varies remarkably even when the distance is fixed. However when the link quality is averaged it falls smoothly as distance increases.

Data plotted in Figure 2.4(a) is used to derive an empirical link quality model by calculating the mean and the standard deviation of the data assuming that the data follows the normal distribution.⁶ Later this empirical model is used to obtain the radio coverage area, shown in Figure 2.4(b), which is certainly not a perfect disk as the path loss model implies.

In short, this empirical study can be seen as support for the log-normal model.

⁶In (Hekmat, 2006a), there are *limited* measurements which validates this assumption.

Chapter 3

Topology Control

This chapter introduces topology control. In §.3.1 three definitions of topology control are presented and discussed. The primary task of topology control techniques is to construct *nice* network topologies over which routing protocols and distributed algorithms run with low cost. These nice topologies have certain properties which are presented in §.3.2. Topology control techniques can be classified into certain distinct groups. One classification is presented in §.3.3. Topology control algorithms have two objectives: saving energy and reducing *interference*. The argument which shows the former objective is achieved through topology control algorithms is clear. Sadly, regarding the latter objective the argument is rather vague and hence §.3.4 is provided to give insight into this issue.

3.1 Definition of Topology Control

A *network topology* is a structure of nodes and links (Buchin and Buchin, 2007) of a graph embedded in \mathbb{R}^d where $d = 1, 2, 3$. Recall that in this thesis we consider the case where $d = 2$.

Topology control techniques applied to a given network, modelled as a graph $G = (V, E)$, result in producing a subgraph, $G' = (V, E')$, which meets desired properties. Although it is well studied, there is no standard definition of topology control.

In (Labrador and Wightman, 2009) topology control is defined as an iterative process consisting of two phases: the construction phase and the maintenance phase. Initially just after the set-up of a network, every node transmits with its maximum power and identifies its neighbour. Then a construction phase has to be initialized and a topology control algorithm is run over the network to compute a *reduced topology* (a subgraph) such that certain network-wide properties are satisfied. The dynamic nature of wireless *ad hoc* networks can change the topology constructed in the construction phase causing the network to lose some of the required properties such as connectivity. There-

fore another algorithm has to monitor the state of the constructed topology and triggers the construction phase when required. This is the maintenance phase.

This definition is quite general and not exclusive. However it points out a serious respect of topology control, the maintenance, which is usually ignored in the literature. A topology control algorithm may not be costly by itself but if it is not robust to dynamism it will be over executed causing the network energy to drop sharply. In this work we have chosen to follow the main stream by separating the construction phase from the maintenance phase and only focusing on the construction phase.

In (Santi, 2005c) the topology control is defined as “the art of coordinating nodes’ decisions regarding their transmitting ranges in order to generate a network with desired properties while reducing node energy consumption and/or increasing network capacity”.

It is obvious that this definition excludes any technique that alters the network topology and produces a reduced topology as long as it does not change transmitting ranges. More specifically this definition excludes techniques used to construct a *virtual backbone*¹ by computing a maximal independent set, an approximation to a minimum dominating set, etc. since the common assumption thereof is that every node transmits with the same power. However these techniques can be considered as topology control techniques based on the definition of topology control given in (Wiki04, 2013) which defines it as a technique used to alter the network topology to get a new topology, over which distributed algorithms run with low cost. The definition of (Santi, 2005c) is the closest to our view.

Based on Eq.(2.4.2) the cost of a wireless link between two nodes grows as a power law of the distance separating them which implies that nodes can save power by discarding long links and communicating only using short links. The topology control algorithm is supposed to assign to each node a transmitting range which results in removing long links and the whole reduced topology meets particular properties mentioned in the following section.

3.2 Criteria of Good Topology

Let $G = (V, E)$ be a max-power graph. The task of a topology control algorithm is to produce a subgraph, $G' = (V, E')$, such that some of the following properties are met.

- *Connectivity* is stressed and considered as a basic requirement that G' has to satisfy. Connectivity implies that if there is a path connecting u and v in G , a path connecting them must exist in G' as well. If G' is fragmented it might paralyse the flow of information. In addition,

¹See (Oliveira and Pardalos, 2011, ch. 11).

from a practical point of view connectivity guarantees that every user of a particular network will be provided with the applications offered by the network. In certain scenarios dis-connectivity has devastating consequences; for example a group of soldiers in a battlefield.

- *Power spanner*: based on Eq.(2.4.2) the cost of a wireless link between two nodes u and v is given by (Rodoplu and Meng, 1999)

$$P(u, v) = \mathcal{P}|uv|^\alpha, \quad (3.2.1)$$

where \mathcal{P} is the predetection threshold, α is the path loss exponent and $|uv|$ denotes the Euclidean distance (L^2 -norm) between u and v here and henceforth. The cost of a path is the sum of the costs of all its constituent links.

A graph G' is called a power spanner if it has a constant *power stretch factor*. The power stretch factor is defined as (Buchin and Buchin, 2007)

$$c = \max_{u,v \in V} \frac{P_{G'}(\Pi(u, v))}{P_G(\Pi(u, v))}, \quad (3.2.2)$$

where $P_{G'}(\Pi(u, v))$ and $P_G(\Pi(u, v))$ denote the cost of the minimum power path connecting u and v in G' and G respectively. Sometimes G' is referred to as a c -power spanner.

By now it should be clear that power spanner is a stronger property than connectivity in a sense that a power spanner does not only require any pair of nodes to be connected by a path in G' if they are connected by a path in G but also it requires that the cost of the minimum power path connecting the pair of nodes in G' to be at most a given constant times the cost of the minimum power path in G connecting the same pair of nodes.

Power spanner is an important property. As we know that nodes prefer to use short links and discard long ones to save their own energy and if the spanner property is not ensured too many long links might be removed and the energy consumed by paths in G' might grow unacceptably high.

- *Sparseness*: when many long links are removed, the resulting graph will be sparse. It is clear that there is a trade-off between sparseness and connectivity since if too many links are removed the reduced topology will be disconnected.
- *Bounded degree*: a graph has a bounded degree if the maximum node degree is bounded by a constant. Bounded degree is a stronger property than sparseness, i.e., there are classes of graphs that are sparse but their maximum node degree is $O(n)$; for example RNG (Buchin and Buchin, 2007). The bounded degree property reduces node overhead and makes

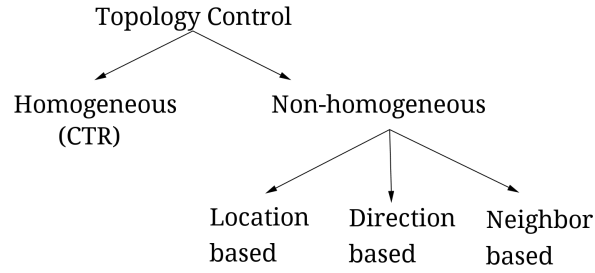


Figure 3.1: Taxonomy of topology control techniques.

the topology robust to dynamism (adaptability) because if the topological information of some nodes has changed only a few nodes have to deal with that.

It is claimed that low interference is a direct consequence of sparseness and the bounded degree property; see for example (Wattenhofer *et al.*, 2001). The validity of this claim is discussed in §.3.4.

- *Symmetry*: we assume that links are symmetric. This is a common requirement of the reduced topology. The reason for considering only symmetric links is that many topology control algorithms, routing protocols and media access control (MAC) protocols assume links are bidirectional as they require messages to be acknowledged (Buchin and Buchin, 2007; Santi, 2005b).

3.3 Taxonomy of Topology Control

Based on the definition of topology control of (Santi, 2005c), topology control techniques can be classified as shown in Figure 3.1.

Generally topology control techniques can be grouped into two categories (Santi, 2005c): homogeneous and non-homogeneous. In homogeneous topology control the problem of interest is to find the CTR; see §.1.4. There is a variant of this problem in which the transmitting range is fixed and the problem is to find the minimum node density that ensures connectivity. This problem is solved in (Bettstetter and Hartmann, 2005) when the network is modelled by the log-normal model.

The task of non-homogeneous techniques is to assign to each node a different transmitting range such that certain network-wide properties are satisfied. Properties of interest include connectivity, spanner, sparseness, etc. Based on the information they require to build the reduced topology, the non-homogeneous techniques are divided into three categories; see Figure 3.1. Since the location-based techniques are not robust to the dynamism and expensive in terms of the information they require to build the topology we decide to

ignore them. The pioneer work in direction-based techniques is the CBTC algorithm which is studied in details in §.4.2. The neighbour-based techniques is the most suitable for wireless *ad hoc* networks. A typical example of these techniques is the XTC algorithm introduced in (Wattenhofer and Zollinger, 2004) and discussed in §.4.1.

3.4 Topology Control and Interference

Network nodes use electromagnetic waves of the *same* frequency when communicating with each other. As a result multiple parallel transmissions can *interfere*. If two signals that are meant to be received by the same node interfere, the node may either receive none of them or receive only the stronger signal. Thus interference reduces network capacity as well as increasing the number of retransmissions which occur when signals are not correctly received. A topology control algorithm is, then, expected to confine interference to increase network capacity and prolong network lifetime by saving energy which could have been wasted as a result of retransmission if interference is not restricted.

Interference might occur in one of the following scenarios:

- *Hidden terminals* which is depicted in Figure 3.2(a). As the figure shows, u and v are out of the vicinity of each other and hence they cannot detect each other. As a result they might transmit simultaneously to w causing interference to take place at w . Node w might receive nothing or receive only the stronger signal. A solution to this problem is provided by MAC (Volbert, 2005), (Santi, 2005c). The basic idea of MAC is that it prevents any communication from taking place if it can interfere with an ongoing transmission. This prevention process is done through exchanging control messages between nodes willing to transmit to one another and keeping a network allocation vector which traces current transmissions.
- As a result of using MAC any node within the vicinity of either u or v cannot transmit while the transmission from u to v is taking place even if the destination sink is out of the vicinity of both u and v . This scenario is referred to as *exposed terminals* which results in decreasing the network capacity. Figure 3.2(b) illustrates this scenario. Although node z is out of u 's vicinity, node w cannot transmit to z unless the current transmission between u and v is finished.
- Since we allow individual nodes to set their transmitting ranges, a third type of interference might occur which is called *asymmetric interference* and depicted in Figure 3.2(c). As the figure shows, node w has a very short transmitting range so its current transmission to node z cannot be

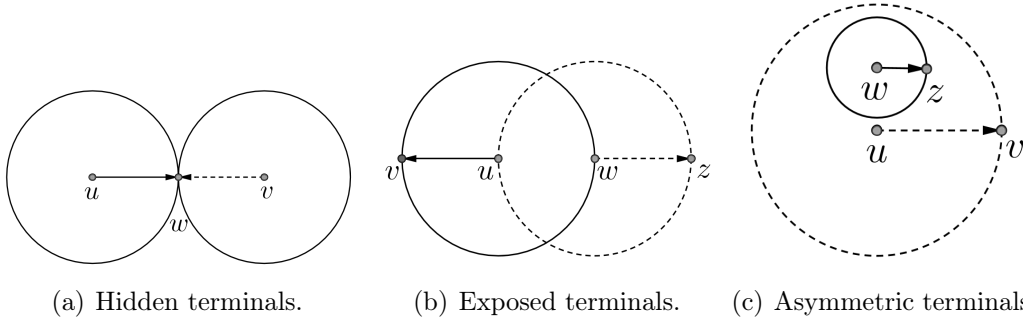


Figure 3.2: Different scenarios of interference. A solid arrow represents the ongoing transmission. A dashed arrow represents potential interfering transmissions.

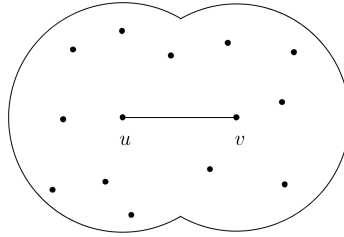


Figure 3.3: Link interference. The number of nodes inside the two disks is the interference of the link (u, v) .

detected by node u which will cause interference if it sends to v while the transmission between w and z is still taking place.

3.4.1 Interference Model

The interference of a link $e = (u, v)$ is defined (Burkhart *et al.*, 2004) as the number of nodes covered by the two disks centred at u and v with radius $|uv|$. Formally speaking, let $D(u, r)$ denote the disk centred at u with radius r . The coverage of the link e is defined as (Burkhart *et al.*, 2004)

$$\text{Cov}(e) = \{w \in V \mid w \in D(u, |uv|) \vee w \in D(v, |uv|)\}.$$

The interference of the link e is, then, defined as

$$I(e) = |\text{Cov}(e)|. \quad (3.4.1)$$

Figure 3.3 illustrates the definition of link interference. Simply the interference of a link is the cardinality of the coverage set of that link.

The interference of a graph G is the maximum link interference which can be expressed as

$$I(G) = \max_{e \in E} I(e) \quad (3.4.2)$$

It is clear that every node in $\text{Cov}(e)$ will be affected by the transmission over e . In (Volbert, 2005), however, the interference is defined based on the effect experienced by the link e as a result of the transmission over other links. Mathematically, the interference of a link $e = (u, v)$ is defined as (Volbert, 2005)

$$I(e) = |\{e' \in E \setminus \{e\} | u \in D(e') \vee v \in D(e')\}|, \quad (3.4.3)$$

where $D(e)$ denotes the disk induced by the link e as defined above.

In (Moaveni-Nejad and Li, 2005) three models of interference are given. The first model is based on interference caused by the transmission over a link which is the same as the interference model defined in (Burkhart *et al.*, 2004) and given by Eq.(3.4.1). The other two models are based on interference caused when a node transmits. The first node-based interference model defines interference as the maximum interference of all links incident to a given node where link interference is as defined by Eq.(3.4.1). The second model defines the interference of a node as the number of nodes within its vicinity.

All these definitions of interference are solely based on the network topology not on the network traffic which depends on many factors including the type of applications provided by the network.

In what follows the unqualified term “interference” refers to the interference model provided in (Burkhart *et al.*, 2004) and given by Eq.(3.4.1) in the context of link interference and by Eq.(3.4.2) in the context of graph interference.

3.4.2 Interference Reduction

Before (Burkhart *et al.*, 2004) all proposed topology control algorithms that attempt to lower interference claim that low interference is a direct consequence of sparseness and the bound degree property without providing a clear argument. All these algorithms remove long links and keep short ones using different ways of achieving this goal. However sometimes a short link may have a higher interference compared with a longer link which implies that in some situations a node has to decide either to keep a longer link which lowers interference (network-wide view) or to discard it and keep a shorter link and save its own power (node-wide view) but not both.

As proved in (Burkhart *et al.*, 2004) there are instances of graphs which have $\Omega(n)$ interference if the RNG is included as a subgraph. Unfortunately all topology control algorithms prior to (Burkhart *et al.*, 2004) include RNG as a subgraph; see Theorem 3.1.

It is worth noting that there are instances in which the optimum topology does not have a constant interference if links are symmetric and connectivity is to be maintained. For example the graph shown in Figure 3.4 which shows a chain of nodes that are exponentially separated where the distance between nodes i and $i + 1$ is 2^i . If links are symmetric and connectivity is to be maintained, the interference of this topology is $O(n)$ because all nodes except

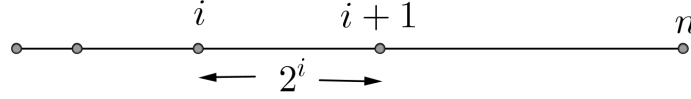


Figure 3.4: Optimum topology with non constant interference.

the left most one establishing a link with its right neighbour will cover all nodes on its left side. Therefore the quality of a given topology in terms of interference has to be determined compared with the interference-optimal topology of the same network.

Theorem 3.1. (*Burkhardt et al., 2004*). *There exist classes of graphs for which any topology control algorithm establishing only symmetric links and including RNG as a subgraph will produce $\Omega(n)$ interference while the interference of the optimum topology is constant.*

Proof. Assume that nodes can increase their transmitting ranges arbitrarily. Consider the network depicted in Figure 3.5(a) which is an extension of the network shown in Figure 3.4. Each horizontal node h_i has a corresponding vertical node v_i placed at a distance equal to d_i where d_i satisfies $d_i > 2^{i-1}$ which implies that the distance from any horizontal node to its corresponding vertical node is larger than the distance to its left neighbour. Between any two vertical nodes v_{i-1}, v_i , a helper node t_i is placed such that $|h_i t_i| > |h_i v_i|$ and $|h_{i-1} t_i| > |h_{i-1} v_{i-1}|$ hold. The RNG of this network is depicted in Figure 3.5(b). Note that almost one third of the nodes belong to the horizontal chain which pushes the interference of RNG up to $\Omega(n)$. An interference-optimal topology is depicted in Figure 3.5(c) which has a constant interference. \square

As we mentioned in §.2.1, topology control algorithms are desired to be local. Unfortunately as Theorem 3.2 shows there are classes of graphs for which no local algorithm can produce non trivial approximation of the optimum topology if connectivity is to be maintained.

Theorem 3.2. (*Burkhardt et al., 2004*). *There exist classes of graphs for which any local algorithm required to construct connected reduced topology will have unbounded interference.*

Proof. Consider Figure 3.6 where we have n nodes. Each node has a maximum transmitting range $|uv|$. Assume that the shaded area contains $\Omega(n)$ nodes connected with $O(1)$ interference and each node w in the shaded area satisfies $|wv| < |uv|$ and $|wu| > |uv|$. It is obvious that the link (u, v) has $\Omega(n)$ interference. In addition, suppose that there is a path (the dashed line) connecting u and v with $O(1)$ interference. Based on this setting any local algorithm is forced to include the link (u, v) if it maintains connectivity since

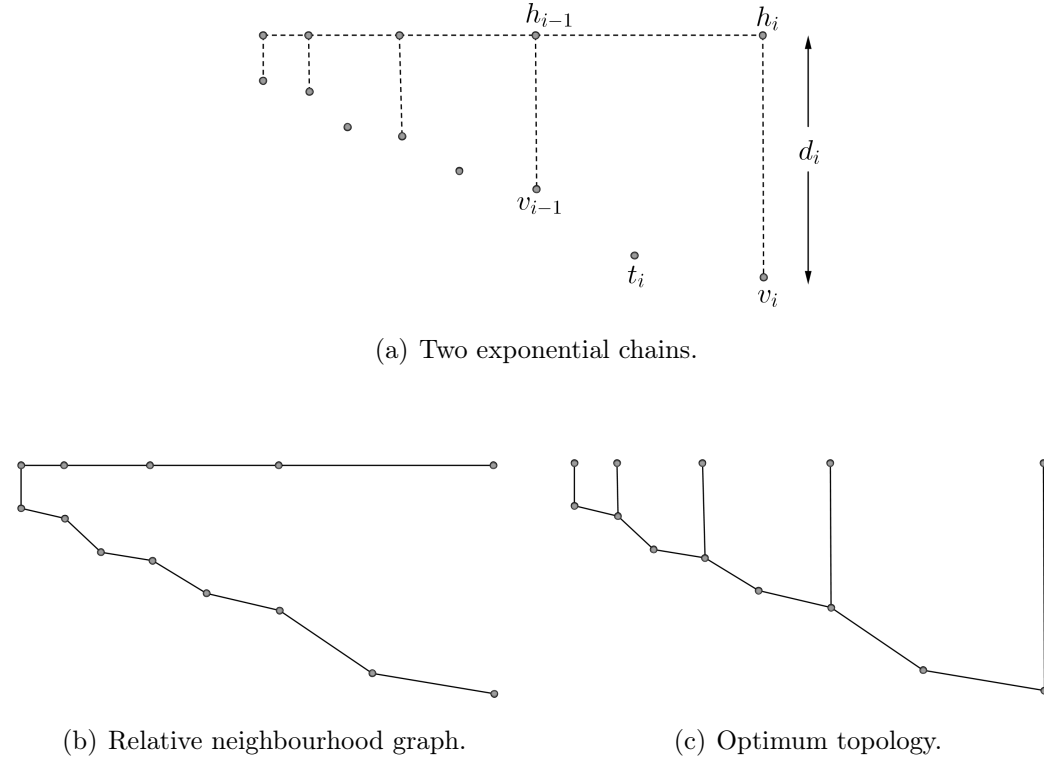


Figure 3.5: Interference.

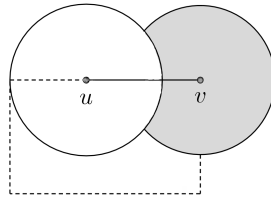


Figure 3.6: The shaded area contains $\Omega(n)$ nodes connected with $O(1)$ interference. The dashed line represents a path connecting u and v with $O(1)$ interference.

it does not know that there is a path connecting u and v . Consequently interference will be pushed up to $\Omega(n)$ while the optimum interference topology has $O(1)$ interference. \square

Note that the two theorems above assume that nodes are not distributed uniformly at random. However almost all topology control studies assume uniform spatial distribution. In addition, they do not imply that by producing a sparse graph interference remains as the same as the interference of the max-power graph. Many studies show that a significant reduction in interference is achieved by controlling the transmitting range; see for example (Song *et al.*, 2005) and §.4.3.

Chapter 4

Non-homogeneous Topology Control

4.1 XTC

4.1.1 Introduction

The XTC algorithm (Wattenhofer and Zollinger, 2004) is suitable for wireless *ad hoc* networks because it is strictly local as well as very cheap in terms of the information it requires to compute the reduced topology. It does not require either position or direction information, rather it requires only each node to be able to detect its neighbours and order them according to a certain criterion. Regarding the name of the algorithm, “TC” obviously stands for *topology control*. The letter “X” is controversial even among the authors. Suggestions include *exotic*, *extreme*, *exceptional* and *exemplary*. Unfortunately none of these suggestions make perfect sense to this reader!

4.1.2 The Algorithm Description

When it is run on a network modelled as a random geometric graph, the algorithm computes a variant of a relative neighbourhood graph locally and it computes an exact relative neighbourhood graph if ties are broken in the latter one.

The algorithm is very simple and consists of three main phases: a) neighbour ordering, b) order exchanging and c) link selection.

Nodes are required to order their neighbours totally according to a certain metric for example a node can order its neighbours based on packet reception rate, the strength of the signal received from its neighbours, etc. Let \prec_u ¹ denote the totally ordered set established by node u . The notation $v \prec_u w$

¹We use this notation in a very sloppy way as used in (Wattenhofer and Zollinger, 2004) since we use it to refer to the totally ordered set *as well as* to the binary operation itself, i.e., the total order.

means node v appears before node w in the totally ordered set of node u . Informally speaking, it means v is a “better” neighbour of u than w .

After ordering its neighbours, node u broadcasts its totally ordered set and receives the totally ordered sets from its neighbours.

Upon receiving the totally ordered sets, node u starts the link selection phase by considering its neighbours in descending order, i.e., “good” neighbours are considered first. Node u establishes a direct link with a neighbour v if there is no a better neighbour of u which is reachable from v more easily than from u itself. In other words node u establishes a direct link with v if and only if they are relatively “best” or relatively “closest” to each other or if and only if they do not have a common neighbour which is “closer” to any one of them than the other.

The above description of the algorithm considers the view of node u and when every node does exactly what u does the resulting graph is referred to as $G_{XTC} = (V, E_{XTC})$ where E_{XTC} is defined as

$$E_{XTC} = \{(u, v) \in V \times V \mid v \in N_u, \forall u \in V\}.$$

Algorithm 1 is a formal description of XTC from the point of view of node u .

Algorithm 1 XTC from the viewpoint of node u .

```

Compute  $\prec_u$ .
Broadcast  $\prec_u$  and receive the totally ordered sets from all neighbours.
 $N_u = \{\}$  # the final set of neighbours of  $u$ .
 $\tilde{N}_u = \{\}$  # the set of abandoned neighbours.
while  $\prec_u \neq \{\}$  do
     $v \leftarrow$  the best neighbour in  $\prec_u$ .
    if  $\exists w \in N_u \cup \tilde{N}_u : w \prec_v u$  then
         $\tilde{N}_u \leftarrow \tilde{N}_u \cup \{v\}$ 
    else
         $N_u \leftarrow N_u \cup \{v\}$ 
    end if
     $\prec_u \leftarrow \prec_u \setminus \{v\}$ 
end while
    
```

Figure 4.1 shows the result of implementing XTC on a network of 100 nodes modelled as a random geometric graph. It is clear that while preserving connectivity G_{XTC} is very sparse. XTC remarkably confines interference as we shall see in §.4.3. In this particular case the original topology has interference equal to 46 while the G_{XTC} has interference only equal to 7. As we will see, generally G_{XTC} is not a spanner, yet in practice it is very unlikely that the cost of the shortest path between a given pair of nodes in G_{XTC} is $O(n)$. For example in this particular case the G_{XTC} has a length stretch factor equal to 7 implying that the power stretch factor is bounded by 7^α using Lemma 2 in (Li

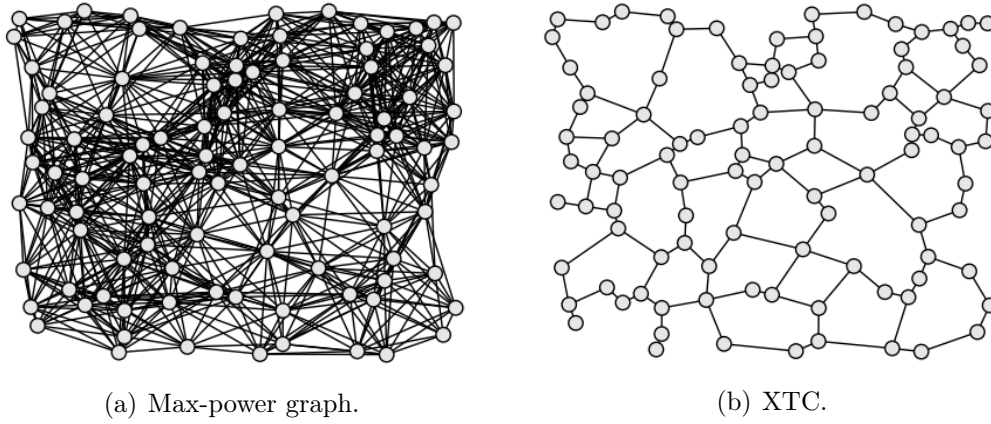


Figure 4.1: XTC applied to a network.

et al., 2001). Generally the performance of G_{XTC} is expected to be much better than the performance of the max-power graph but we have not considered investigating that as we are only interested in the mathematical properties of the topology produced by XTC irrespective of the practical aspects which can be studied through network simulators.

We have seen in §.3.2, it is desirable to have only symmetric links in the reduced topology. Fortunately, as the following theorem shows, links in G_{XTC} are symmetric.

Theorem 4.1. (*Wattenhofer and Zollinger, 2004*). *If $v \in N_u$, then $u \in N_v$ and vice versa.*

Proof. For the sake of contradiction assume that $v \in N_u$ while $u \notin N_v$. The assumption $u \notin N_v$ implies that there exists $w \in N_v \cup \tilde{N}_v$ such that $w \prec_u v$. The fact that w is already in $N_v \cup \tilde{N}_v$ when u is processed implies $w \prec_v u$. Since $w \prec_u v$, by the time when u is considering the inclusion of v we have $w \in N_u \cup \tilde{N}_u$. Together with $w \prec_v u$ result in including v in \tilde{N}_u which contradicts the assumption that $v \in N_u$. \square

4.1.3 Properties of the topology produced by XTC

The algorithm works successfully in general weighted graphs. However in the specific case when the network is modelled by a random geometric graph, the reduced topology is proved to satisfy many interesting properties. We prove by mathematical induction that the graph G_{XTC} preserves the connectivity as Theorem 4.2 shows. In (Wattenhofer and Zollinger, 2004) connectivity is proved by contradiction. In addition, G_{XTC} is proved to have a bounded degree as Theorem 4.3 shows and to be planar as Theorem 4.4 shows.

Theorem 4.2. *Let $G = (V, E)$ be a random geometric graph and $G_{XTC} = (V, E_{XTC})$ be the graph obtained after running XTC on G . Then G_{XTC} is*

connected if and only if G is connected.

Proof. If G_{XTC} is connected then so is G since $G_{XTC} \subset G$.

We prove the opposite direction by mathematical induction. To do that we order links of G_{XTC} according to their length and we define a rank of a link (u, v) as follows:

If (u, v) is among the shortest links in the graph then its rank is 1. Otherwise, the rank is the smallest integer k with the condition that for any link (u', v') with $|u'v'| < |uv|$, the rank of (u', v') is less than k .

For the base case if $(u, v) \in E$ is the shortest link it follows easily that $(u, v) \in E_{XTC}$ since if $(u, v) \notin E_{XTC}$ there exists $w \in N_u \cup \tilde{N}_u$ such that $w \prec_v u$. By definition $(w, v) \in E$ but that contradicts the assumption (u, v) is the shortest link in E .

For the induction hypothesis assume that every link in G of rank at most $k - 1$ is connected through a path in G_{XTC} .

Now we wish to prove that a link of rank k is also connected via a path in G_{XTC} . Consider a non-connected pair of nodes u and v whose direct link in G is of rank k . Since $(u, v) \notin E_{XTC}$ there exists $w \in N_u \cup \tilde{N}_u$ such that $w \prec_v u$. From $w \in N_u \cup \tilde{N}_u$ we have $w \prec_u v$. Since $(u, v) \in E$ then by definition $(u, w), (v, w) \in E$ as well. Note that (u, w) and (v, w) are links in G whose rank is at most $k - 1$. Therefore from the induction hypothesis, paths from u to w and from v to w exist in G_{XTC} . Therefore a path from u to v exists in G_{XTC} since links in G_{XTC} are symmetric by Theorem 4.1. \square

Although it preserves connectivity, unfortunately G_{XTC} is not guaranteed to be a power spanner which means that the power stretch factor of G_{XTC} can be as high as $O(n)$ times the power stretch factor of G . The construction shown in Figure 4.2 together with Lemma 4.1 are used (Li *et al.*, 2001) to show that G_{XTC} has a power stretch factor equal to $n - 1$. Let c denotes the power stretch factor of G_{XTC} . We show that $n - 1 - \epsilon < c \leq n - 1$ for any positive ϵ which implies $c = n - 1$.

Lemma 4.1. (Li *et al.*, 2001) *Let H be a subgraph of G . The graph H has a power stretch factor c if and only if for any link (u, v) in G but not in H , $P_H(\Pi(u, v)) \leq c|uv|^\alpha$ where $P_H(\Pi(u, v))$ denotes the cost of the minimum power path connecting u and v in H and α is the path loss exponent.*

Firstly if (u, v) is in G , in the worst case, a path connecting u and v in the Euclidean minimum spanning tree (EMST) consists of at most $n - 1$ links each of them at most $|uv|$. The cost of such a path is at most $(n - 1)|uv|^\alpha$. By using Lemma 4.1, the power stretch factor of EMST is at most $n - 1$. Since EMST is a subgraph of G_{XTC} we have $c \leq n - 1$.

Secondly consider the construction shown in Figure 4.2. In the figure two cases are considered: when n is even, say $n = 2m$ for a positive integer m which is shown in Figure 4.2(a) and when it is odd which is shown in Figure 4.2(b). In

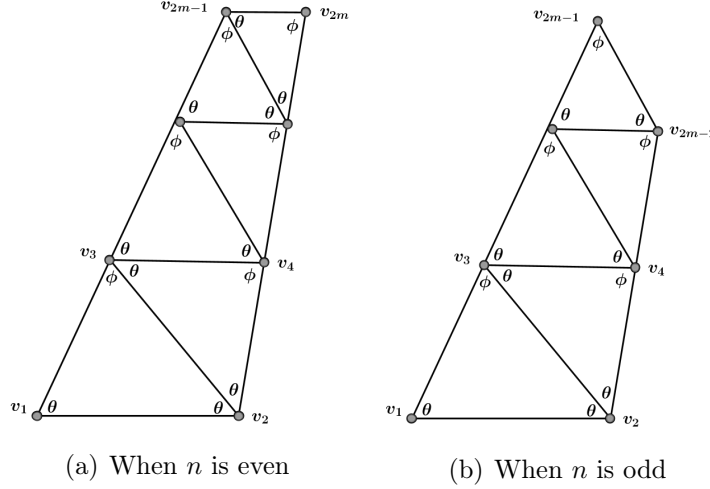


Figure 4.2: Non constant stretch factor of G_{XTC} .

the figure $\phi = \frac{\pi}{3} + 2\delta$ and $\theta = \frac{\pi}{3} - \delta$. Given this construction, it is easy to check that the path connecting v_1 and v_2 in G_{XTC} is $v_1, v_3, \dots, v_{2m-1}, v_{2m}, \dots, v_4, v_2$. When $\delta \rightarrow 0$ the length of each link in the path tends to $|v_1 v_2|$. Therefore

$$c = \frac{P_{G_{XTC}}(\Pi(v_1, v_2))}{P_G(\Pi(v_1, v_2))} \longrightarrow n - 1,$$

where $P_{G_{XTC}}(\Pi(v_1, v_2))$ and $P_G(\Pi(v_1, v_2))$ are the costs of the minimum power path connecting v_1 and v_2 in G_{XTC} and G respectively.

Since δ is an arbitrary positive number, we can find a sufficiently small positive number ϵ such that $c > n - 1 - \epsilon$. When n is odd the above argument is still valid. We conclude that $c = n - 1$ as claimed.

Theorem 4.3. (*Wattenhofer and Zollinger, 2004*). *The maximum degree in G_{XTC} is 6.*

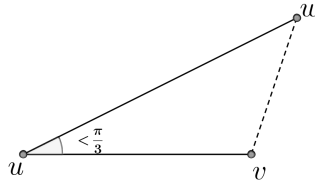


Figure 4.3: A triangle with $\angle wuv < \frac{\pi}{3}$.

Proof. The claim is proved by showing that any two adjacent links in G_{XTC} form an angle which is at least $\frac{\pi}{3}$. Suppose that u has two neighbours v and w that enclose an angle less than $\frac{\pi}{3}$; see Figure 4.3. Assume that $|uv| < |uw|$.

Since $\angle wuv < \frac{\pi}{3}$ and $|uv| < |uw|$ it follows that $|vw| < |uw|$. Therefore when u decided on the inclusion of w , node v has already been processed, i.e., $v \in N_u \cup \tilde{N}_u$. Together with $v \prec_w u$ (since $|vw| < |uw|$) result in including w in \tilde{N}_u which contradicts the assumption that $w \in N_u$. Therefore $\angle wuv \geq \frac{\pi}{3}$ and the theorem follows. \square

Theorem 4.4. (Wattenhofer and Zollinger, 2004). G_{XTC} is planar.

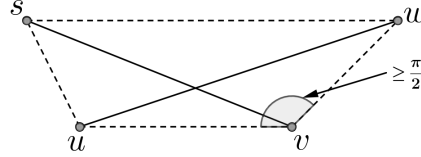


Figure 4.4: Links (u, w) and (v, s) cross and $\angle uvw \geq \frac{\pi}{2}$.

Proof. Suppose that G_{XTC} is not planar which means that at least there exist two crossing links. Let these two links be (u, w) and (v, s) of the quadrilateral $uvws$ shown in Figure 4.4. One angle of the quadrilateral $uvws$ must be at least $\frac{\pi}{2}$. Without loss of generality, let this angle be $\angle uvw \geq \frac{\pi}{2}$. Therefore we have $|uv| < |uw|$ and $|vw| < |uw|$. These two conditions lead to $w \in \tilde{N}_u$ which contradicts $w \in N_u$. \square

4.2 Cone-Based Topology Control Algorithm

4.2.1 Introduction

The cone-based topology control algorithm (CBTC) is firstly presented in (Wattenhofer *et al.*, 2001) and improved in (Li *et al.*, 2005) by adding a set of optimizations to the original algorithm and evaluating its performance using the NS-2 network simulator. Although CBTC is more expensive than XTC in terms of the information required to compute the reduced topology it is still cheap as it only requires each node to determine the relative directions of its neighbours. Generally direction information is considered cheaper than position information because to obtain direction information nodes are not required to be equipped with GPS. Instead they are equipped with directional antennas and from measuring the time difference of the received signal at each antenna, the direction of the received signal can be estimated.

4.2.2 The Algorithm Description

The algorithm computes a variant of a Yao graph when the network is modelled by a random geometric graph.

The algorithm reduces the transmitting range of each node by assigning to each node the minimum transmitting range that ensures that in every cone of angle θ , where θ is a parameter passed to the algorithm, around the node there exists at least one neighbour unless the cone is empty. The algorithm is distributed and runs in each node separately either synchronously or asynchronously. The algorithm can be informally described from the viewpoint of a particular node, say u , as follows: node u initially sets its transmitting range to the minimum value and broadcasts with this range, then it collects acknowledgements from the discovered neighbours. Upon the discovery of a neighbour, node u computes the relative direction of the newly discovered neighbour and builds a cone of angle θ around itself having the line \overline{uv} as a bisector where v is the newly discovered neighbour; see Figure 4.5. Then it checks whether or not there is a gap of angle θ around itself. If there is, it increases its transmitting range and repeats the same process. It stops running the algorithm if one of the following two conditions satisfies first: a) if there is no gap of angle at least θ . b) if the transmitting range hits its maximum value.

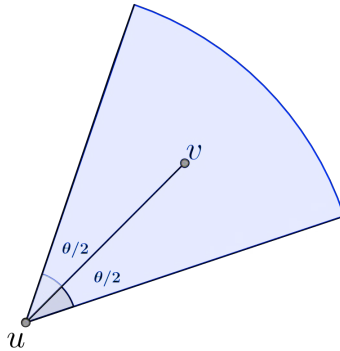


Figure 4.5: A cone of angle θ around u having \overline{uv} as a bisector.

If there is a cone of angle at least θ which contains no neighbour of u the algorithm terminates due to the second condition which implies that the transmitting range of u will be set to the maximum value and that disagrees with the very purpose of the algorithm as it is supposed to reduce the transmitting range at each node. To resolve this problem an optimization called *shrink-back operation* has been proposed in (Li *et al.*, 2005) which is an additional phase that only nodes with the maximum transmitting range are required to run to reset their transmitting range such that their coverage does not change, where

the coverage is defined as (Li *et al.*, 2005):

$$\{\phi : |\phi - \varphi| \bmod 2\pi \leq \theta/2, \text{ for some } \varphi \text{ in the set of discovered directions}\}.$$

The shrink-back operation processes links established by CBTC based on their length in a descending order, i.e, starting with a longest link, and it discards any link whose removal will not create a *new* gap of angle greater than zero. Figure 4.6 explains the shrink-back operation. Suppose that after running CBTC with $\theta = \pi/2$, node u has four neighbours $\{u_1, u_2, u_3, u_4\}$ as shown in the figure. The coverage of u is defined by the shaded area. Then there exists a gap of angle at least θ and hence the transmitting range of u will be set to the maximum value. Consequently node u has to run the shrink-back operation. As a result, the link (u, u_2) will be removed because its removal does not change the coverage. On the contrary, the link (u, u_4) is kept because if it were removed it would have changed the coverage and created a new gap whose angle is larger than zero since the cone with the bisector $\overline{uu_4}$ has part of it *not* covered by all the remaining cones. After running the shrink-back operation, the transmitting range of u will be $|uu_4|$ instead of the maximum value.

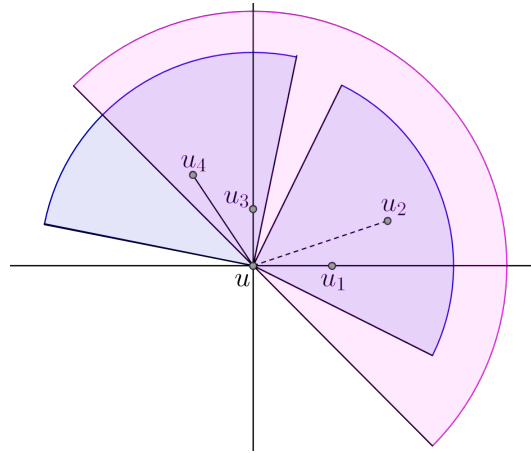


Figure 4.6: Illustration of shrink-back operation.

Algorithm 2 is a formal description of CBTC. In the algorithm description, “Increase” is a function receiving a transmitting range as a parameter and *somehow* increasing it and “Gap” is another function that receives an angle θ and a set of directions D_u as parameters and returns “true” if there is a gap of angle at least θ around u and “false” if there is not.

It is clear that the algorithm computes a variation of a Yao graph; see §.2.1. The difference between the topology computed by CBTC and a Yao graph is that Yao cones are fixed and predetermined whereas CBTC cones are created on-the-fly, i.e., cones are created upon the discovery of neighbours.

Algorithm 2 CBTC from the perspective of node u .

```

 $N_u \leftarrow \{\}$  # the set of discovered neighbours.
 $D_u \leftarrow \{\}$  # the set of directions of discovered neighbours.
 $r_u \leftarrow r_0$  # transmitting range.
while  $r_u \leq r_{\max}$  and  $\text{Gap}(\theta, D_u)$  do
     $r_u \leftarrow \text{Increase}(r_u)$ 
    Broadcast and gather acknowledgements
     $N_u \leftarrow N_u \cup \{v \mid v \text{ is discovered}\}$ 
     $D_u \leftarrow D_u \cup \{d \mid d \text{ is the direction of a newly discovered neighbour}\}$ 
end while
    
```

Because they are created in this way, CBTC cones may overlap while Yao cones cannot. Another difference is that each cone in a Yao graph contains exactly one neighbour while CBTC cones may contain more than one neighbour.

Generally links created by CBTC are not symmetric, i.e., a link (u, v) may exist in the final topology while (v, u) may not. To illustrate this point consider the construction (Li *et al.*, 2005) shown in Figure 4.7. The figure shows a random geometric graph consisting of $\{u, u_1, u_2, u_3, v\}$ with distance threshold r . The nodes are placed such that $|uv| = r$ and angles enclosed by links are as shown in the figure where $0 < \epsilon < \frac{\pi}{12}$. If θ is chosen such that $\frac{2\pi}{3} < \theta \leq \frac{5\pi}{6}$ then after running CBTC we have $N_u = \{u_1, u_2, u_3\}$ and $N_v = \{u\}$ which means that (v, u) exists in the final topology while (u, v) does not. The choice $\theta > \frac{2\pi}{3}$ is not an accident as it has been proved in (Li *et al.*, 2005) that when $\theta \leq \frac{2\pi}{3}$ asymmetric links can be removed without harming the connectivity of the reduced topology.

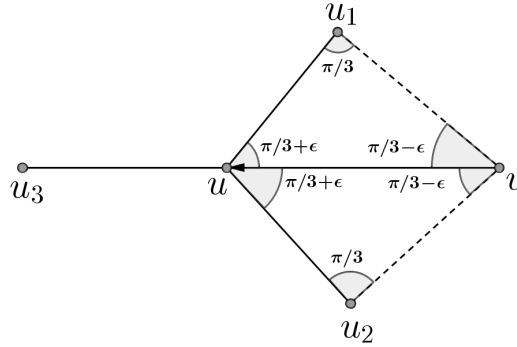


Figure 4.7: While (v, u) is in the reduced topology, (u, v) is not.

In the next section we are going to study the mathematical properties of the topology produced by CBTC so we need to introduce a notation. Define $N_\theta = \{(u, v) \in V \times V \mid v \in N_u, \forall u \in V\}$ and let $G_\theta = (V, E_\theta)$ where E_θ is the symmetric closure of N_θ ; that is $(u, v) \in E_\theta$ if and only if $(u, v) \in N_\theta$ or

$(v, u) \in N_\theta$. In short, G_θ is the reduced topology computed by CBTC when link directions are ignored.

4.2.3 Properties of the topology computed by CBTC

In this section it is proved that G_θ preserves connectivity if $\theta \leq \frac{5\pi}{6}$. Moreover we prove that if $\theta \leq \frac{\pi}{2}$ then G_θ is a power spanner. Finally we comment on the proof of a power stretch factor obtained for the Yao graph in (Li *et al.*, 2001) to show that their mathematical argument is incorrect.

To prove that G_θ preserves connectivity the following key lemma is used.

Lemma 4.2. (Li *et al.*, 2005) *If $\theta \leq \frac{5\pi}{6}$ and $u, v \in V$ such that $(u, v) \in E$, then either $(u, v) \in E_\theta$ or there exists $u', v' \in V$ such that (a) $|u'v'| < |uv|$ (b) either $u' = u$ or $(u, u') \in E_\theta$ (c) either $v' = v$ or $(v, v') \in E_\theta$.*

Roughly speaking, Lemma 4.2 says that if a link (u, v) in G is not kept in G_θ then one of the situations depicted in Figure 4.8 must occur. The case when $u' = u$ and $v' = v$ is excluded because it is meaningless!

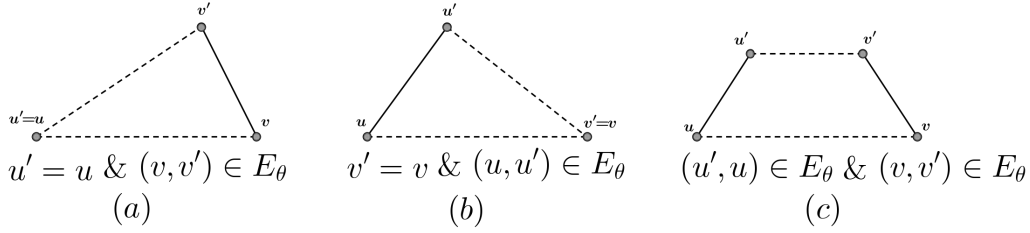


Figure 4.8: Illustration of Lemma 4.2.

Theorem 4.5. (Li *et al.*, 2005). *If $\theta \leq \frac{5\pi}{6}$, G_θ is connected if and only if G is connected.*

Proof. If G_θ is connected then so is G since $G_\theta \subset G$.

The mathematical induction on the link rank where the rank is as defined in Theorem 4.2 is used to prove the opposite direction.

For the base case if (u, v) is the shortest link in G , then according to Lemma 4.2 it must be in G_θ as well.

For the induction hypothesis suppose that for any link in G with rank at most $k - 1$, there is a path in G_θ connecting its end-nodes.

Now consider a link (u, v) of rank k in G . If (u, v) is in G_θ we are done. Otherwise according to Lemma 4.2 there must exist u', v' such that $|u'v'| < |uv|$. Therefore by definition (u', v') must be in G and it is of rank less than k . Then according to the induction hypothesis there exists a path between u' and v' in G_θ which implies the existence of a path between u and v in G_θ since links are symmetric. \square

Theorem 4.6. (*Li et al., 2005*). If $\theta > \frac{5\pi}{6}$, G_θ is not guaranteed to preserve connectivity of G .

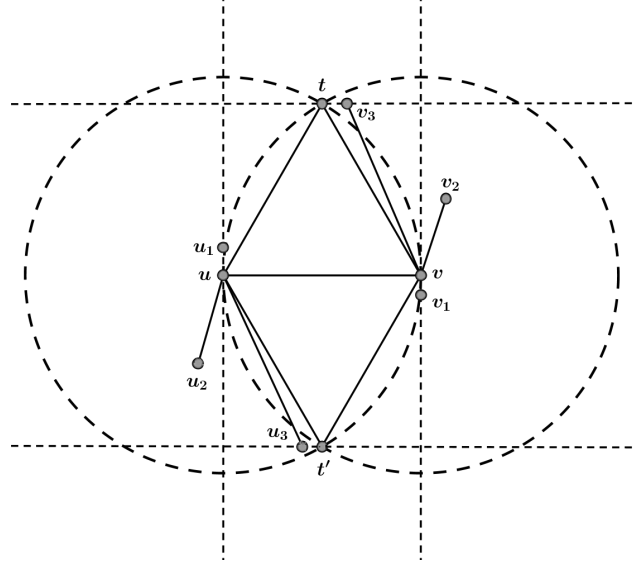


Figure 4.9: The topology G_θ with $\theta > 5\pi/2$ is disconnected.

Proof. Consider the random geometric graph shown in Figure 4.9 where the distance threshold is r . Suppose that $\theta = \frac{5\pi}{6} + \epsilon$ with $\epsilon > 0$. The figure shows two clusters of nodes: u -cluster consisting of $\{u, u_1, u_2, u_3\}$ and v -cluster consisting of $\{v, v_1, v_2, v_3\}$. These two clusters are connected only by the link (u, v) with $|uv| = r$. The points t and t' are the intersection point of the two circles of radius r and centred at u and v . Node u_1 is placed such that $\angle vuu_1 = \frac{\pi}{2}$ and thus $|u_1v| > r$ no matter how small $|uu_1|$ is. Node v_1 is similarly placed but on the opposite side of the line \overline{uv} . Node u_2 is placed such that $\angle u_1uu_2 = \min(\theta, \pi)$. Since $\angle vuu_2 > \frac{\pi}{2}$, $|u_2v| > r$. Node v_2 is placed similarly. The key step in this construction is the choice of nodes u_3 and v_3 . Choose u_3 on the line through t' parallel to the line \overline{uv} such that $\angle u_3uu_1 < \theta$. Since $\theta > \frac{5\pi}{6}$ it is possible to place u_3 in such a position. It is obvious that $|uu_3| < r$ while $|u_3v| > r$. Furthermore we can choose $|vv_1|$ sufficiently small such that $|v_1u_3| > r$. Now it is quite simple to verify that when CBTC is run on this construction with $\theta > \frac{5\pi}{6}$, the set of neighbours of u and v is $\{u_1, u_2, u_3\}$ and $\{v_1, v_2, v_3\}$ respectively. Therefore u -cluster is disconnected from v -cluster and consequently G_θ is disconnected. \square

When $\theta \leq \frac{\pi}{2}$ the topology G_θ does not only preserve connectivity but it also has been proved to be a power spanner. (Wattenhofer *et al.*, 2001) *try* to prove that G_θ has a power stretch factor at most $1 + 2 \sin \frac{\theta}{2}$. Unfortunately as pointed out by (Li *et al.*, 2001) their proof contains some bugs. We obtain the

claimed power stretch factor by avoiding the bugs occurring in (Wattenhofer *et al.*, 2001). Moreover we have proven that this power stretch factor does not *always* hold for G_θ , rather it holds only with probability one. The proof relies on trigonometry and depends on Lemma 4.3 – Lemma 4.6.

Lemma 4.3. *Assume a triangle with angle $\gamma \geq \frac{\pi}{2}$ and sides a, b and c with c opposite to γ . Then $a^k + b^k \leq c^k$ for $k \geq 2$.*

Lemma 4.4. *Assume a triangle with sides a, b and c opposite to angles α, β and γ respectively with $b < a$ and $b < c$ and $\alpha \leq \frac{\pi}{4}$. Then $\gamma \geq \frac{\pi}{2}$ and $a < c$.*

Lemma 4.5. *Assume a triangle with sides a, b and c opposite to angles α, β and γ respectively. Let $a < c$, $b < c$ and $\alpha \leq \frac{\pi}{4}$. Then $a^k + b^k < c^k(1 + 2\sin(\frac{\alpha}{2}))$ for $k \geq 2$.*

Lemma 4.6. *Let $G = (V, E)$ be a random geometric graph, G_θ be as defined above and $\theta \leq \frac{\pi}{2}$. Then after running the algorithm if $(u, v) \in E$ and $(u, v) \notin E_\theta$, then there exists w in the same cone as v with $(u, w) \in E_\theta$ such that $|wv| < |uv|$; see Figure 4.10.*

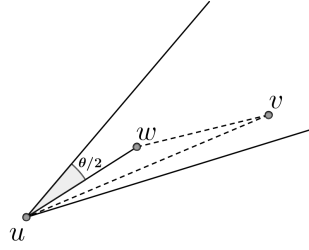


Figure 4.10: Nodes w and v lie in the same cone. $(u, w) \in E_\theta$ and $|wv| < |uv|$.

Proof. Recall that the cone is defined by the angle θ as shown in Figure 4.5. Therefore the angle $\angle wuv \leq \frac{\pi}{4} < \frac{\pi}{3}$, hence $|wv| < \max\{|uw|, |uv|\}$. Note that w is a neighbour of u while v is not. Therefore $|uw| \leq |uv|$ and thus $|wv| < |uv|$. \square

Theorem 4.7. *Let $G = (V, E)$ be a random geometric graph, G_θ be as defined above and $\theta \leq \frac{\pi}{2}$. Then with probability one the topology produced by CBTC is a power spanner with power stretch factor bounded by $1 + 2\sin(\frac{\theta}{2})$.*

Before starting the proof it is worth mentioning that the proof we will present here is incorrect in a particular case which occurs with probability zero and hence we state the theorem in a probabilistic way.

Proof. We wish to prove that for any two nodes u and v in G with minimum power path cost $P_G(\Pi(u, v))$, there exists a path in G_θ with cost bounded from above by $(1 + 2 \sin(\frac{\theta}{2}))P_G(\Pi(u, v))$.

If $(u, v) \in E_\theta$ we are done, i.e., $P_{G_\theta}(\Pi(u, v)) = P_G(\Pi(u, v)) < P_G(\Pi(u, v))(1 + 2 \sin(\frac{\theta}{2}))$.

Suppose that u, v are not connected in G_θ . Then according to Lemma 4.6 there exists a neighbour of u , say u_1 , such that $|u_1v| < |uv|$. Now if u_1 is a neighbour of v , the path is u, u_1, v . Otherwise according to Lemma 4.6 again there exists a neighbour, say u_2 , of u_1 which is in the same cone as v such that $|u_2v| < |u_1v|$. If u_2 is a neighbour of v the path is u, u_1, u_2, v . Otherwise we apply Lemma 4.6 again and so on. Since we are coming closer to v at each step and G_θ is connected, the process has to terminate with u_i for $1 \leq i \leq n$ being a neighbour of v and the path is $u, u_1, u_2, \dots, u_i, v$; see Figure 4.11.

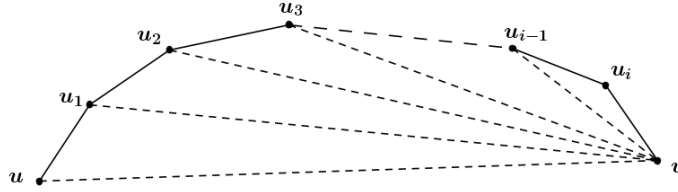


Figure 4.11: A path connecting u and v in G_θ .

The cost of this path is

$$P_{G_\theta}(\Pi(u, v)) = P_{G_\theta}(u, u_1) + P_{G_\theta}(u_1, u_2) + P_{G_\theta}(u_2, u_3) + \dots + P_{G_\theta}(u_i, v).$$

We know that the cost of a direct link (u_i, u_{i+1}) is $P_{G_\theta}(u_i, u_{i+1}) = t|u_i u_{i+1}|^\alpha$. Therefore

$$P_{G_\theta}(\Pi(u, v)) = t|uu_1|^\alpha + t|u_1u_2|^\alpha + t|u_2u_3|^\alpha + \dots + t|u_iv|^\alpha.$$

We assume that all nodes are homogeneous² and therefore the constant t remains the same for all of them. We need to apply Lemma 4.3 repeatedly but before then we have to make sure that it holds for every triangle $u, u_{k-1}u_k$ for $k = 2, 3, \dots, i$; see Figure 4.12.

Since $(u_{k-1}, u_k) \in E_\theta$ we can show that $\angle uu_ku_{k-1} \leq \frac{\pi}{4}$. We have $|u_{k-1}u_k| < |uu_{k-1}|$ and $|u_{k-1}u_k| < |uu_k|$. Then according to Lemma 4.4, $\angle u_ku_{k-1}u \geq \frac{\pi}{2}$ and $|uu_{k-1}| < |uu_k|$. Hence the conditions of Lemma 4.3 are satisfied and therefore

$$|uu_{k-1}|^\alpha + |u_{k-1}u_k|^\alpha \leq |uu_k|^\alpha.$$

²The network devices are the same.

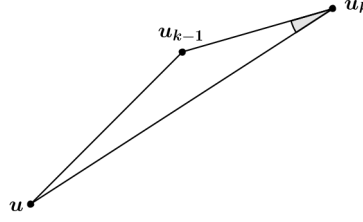


Figure 4.12: A triangle constructed from the path given in Figure 4.11.

Now it is safe to apply Lemma 4.3 repeatedly to have

$$P_{G_\theta}(u, v) \leq t|uu_i|^\alpha + t|u_iv|^\alpha.$$

Let the subscript $k-1$ be i in Figure 4.12. Then Lemma 4.5 can be applied on $\triangle uu_iv$. Therefore

$$P_{G_\theta}(\Pi(u, v)) < t|uv|^\alpha \left(1 + 2 \sin \frac{\theta}{2}\right) = P_G(\Pi(u, v)) \left(1 + 2 \sin \frac{\theta}{2}\right),$$

as desired.

To complete the proof we have to show that if a link (u, v) in G is not preserved in G_θ , then the path connecting u and v in G_θ is *always* as shown in Figure 4.11. This is the case because our proof depends on specific lemmas which hold only for specific structures.

Consider two neighbours u_{k-1} and u_k in the path given in Figure 4.11. According to our choice v and u_k lie in the same cone around u_{k-1} . Then node v can be either above the bisector $\overline{u_{k-1}u_k}$ or below it or on it.

First consider the case where v is below $\overline{u_{k-1}u_k}$ as shown in Figure 4.13. This case is identical to the structure given in Figure 4.11.

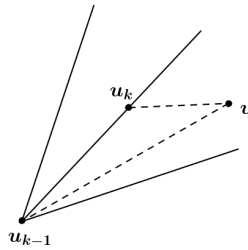


Figure 4.13: The node v is below $\overline{u_{k-1}u_k}$.

Let v be above $\overline{u_{k-1}u_k}$, v' be the mirror image of v with respect to $\overline{u_{k-1}u_k}$ and w be the intersection point between $\overline{u_{k-1}u_k}$ and $\overline{v'v}$ as shown in Figure 4.14. It is easy to show that $\triangle u_k w v$ and $\triangle u_k w v'$ are congruent, hence

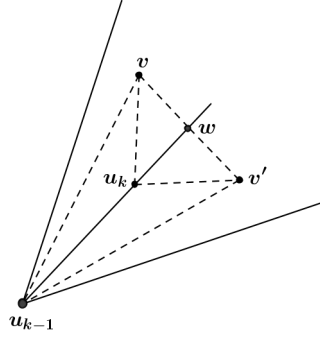


Figure 4.14: The node v is above $\overline{u_{k-1}u_k}$.

$|u_kv| = |u_kv'|$ which can be used to show that the $\triangle u_{k-1}u_kv$ and $\triangle u_{k-1}u_kv'$ are congruent.

Finally if v lies on $\overline{u_{k-1}u_k}$ the above argument does not work. However this case *almost never* happen since the probability that a node lie exactly on $\overline{u_{k-1}u_k}$ is zero. \square

It should be noted that in the original paper (Wattenhofer *et al.*, 2001) it was attempted to prove that $c \leq 1 + 2 \sin(\frac{\theta}{2})$ while we just prove the strict inequality. The equality sign in the original proof makes a significant difference as it implies the obtained result is the *supremum* whereas we do not claim so. Stating that is dangerous as it implies no less upper bound can be found and once a less upper bound is found, the claimed supremum is automatically proven to be false. The literature (Wattenhofer *et al.*, 2001), (Li *et al.*, 2001) and (Bose *et al.*, 2004) that prove the spanner property do not pay much attention to the serious consequence of having the equality sign in the proof of the spanner property. Luckily we note that the equality sign does not hold in all of the aforementioned literature. For instance in (Bose *et al.*, 2004) the equality sign holds only if we have a node with two neighbours lying on the rays that define the cone, see Figure 4.15, which implies that these two neighbours cannot belong to the same cone as each cone is distinguished by only one ray³ implying that their proof does not work in this case as it requires the two neighbours to lie in the same cone. A similar argument is valid regarding (Li *et al.*, 2001).

We have seen that if $\theta > \frac{2\pi}{3}$ connectivity of G_θ is not guaranteed unless we take the symmetric closure of N_θ . Interestingly, when $\theta \leq \frac{2\pi}{3}$ we do *not* have to take the symmetric closure to preserve connectivity and that means asymmetric links can be *removed* from G_θ without violating connectivity (Li *et al.*, 2005). There is a trade-off between transmitting ranges in $G_{\frac{5\pi}{6}}$ and in $G_{\frac{2\pi}{3}}$. In general when $\theta = \frac{5\pi}{6}$ transmitting ranges are lower than transmitting

³If $\theta = \frac{2\pi}{k}$, we can define the cones as $[i, i+1)\theta$ where $i = 0, 1, \dots, k-1$.

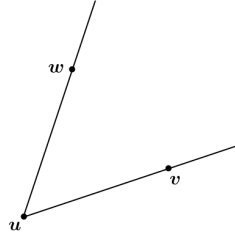


Figure 4.15: Nodes w and v are neighbours of and at the same distance from u .

ranges when $\theta = \frac{2\pi}{3}$, yet if $\theta = \frac{5\pi}{6}$ a node, say u , may use a transmitting range larger than the one set for it by CBTC if u has to reach v where $(v, u) \in N_\theta$ while $(u, v) \notin N_\theta$. It is worth noting that removing asymmetric links remarkably increases the sparseness of G_θ as Figure 4.16 shows which we obtain from implementing CBTC on a network modelled as a random geometric graph.

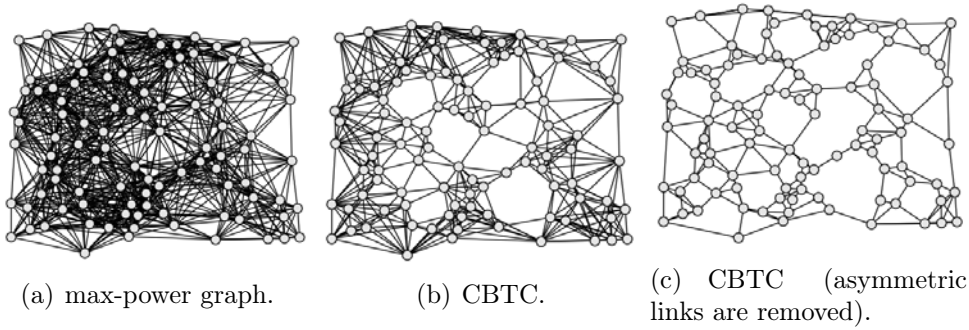


Figure 4.16: CBTC applied to a network with $\theta = \pi/2$.

Even after removing asymmetric links many links can still be removed without violating connectivity. More precisely, a link $(u, w) \in E_\theta$ can be eliminated if there exist links $(u, v), (v, w) \in E_\theta$ such that

$$P(u, v) + P(v, w) \leq q P(u, w), \quad (4.2.1)$$

where $P(u, v)$ is the cost of the link (u, v) and $q \geq 1$. Indeed when $q = 1$ the link (u, w) is an inefficient link, i.e., the cost of direct communication between u and w is higher than the cost of indirect communication between u and w through v . Therefore eliminating such a link is justifiable, however in (Wattenhofer *et al.*, 2001) a link (u, w) is removed even when $q > 1$. Consequently some of the efficient links will be removed and indirect inefficient communications are preferred in this case to direct efficient communications! The reason for this very weird link removal process is the “belief” that by lowering the degree and having as much sparseness as possible interference will be confined. However this conjecture was disproved in (Burkhart *et al.*, 2004) as we have seen in §3.4.

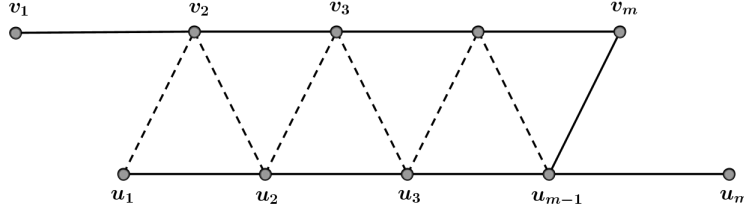


Figure 4.17: All dashed links will be removed according to Eq.(4.2.1).

Another concern is that removing links using Eq.(4.2.1) with $q > 1$, in an uncoordinated manner may harm the power spanner property. (Wattenhofer *et al.*, 2001) conjecture that this link removal process will only increase the energy cost of the minimum power path by only a factor of q and thus will not violate the spanner property. Unfortunately this can be disproved by the following counter example shown in Figure 4.17 which is a random geometric graph with distance threshold r and $V = \{u_i | 1 \leq i \leq m\} \cup \{v_i | 1 \leq i \leq m\}$. The nodes are placed as shown in the figure in such away that $|u_i v_i| = r$, $2 \leq i < m-1$; $|u_i v_{i+1}| = r$, $1 \leq i \leq m-1$; $|v_1 v_2| = |u_{m-1} u_m| = a + \epsilon$ and $|u_i u_{i+1}| = |v_i v_{i+1}| = a$, $1 \leq i \leq m-1$, where $a^\alpha < (q-1)r^\alpha$ and ϵ is chosen such that $a + \epsilon < r$. After running the CBTC with $\theta \leq 2 \times \min(\angle u_{i+1} u_i v_{i+1}, \angle v_i v_{i+1} v_{i+1})$, G_θ will be identical to G . However after running the link removal phase with $q > 1$ all links (u_i, v_{i+1}) , $i = 1, 2, \dots, m-2$ will be removed because

$$\begin{aligned} |u_i u_{i+1}|^\alpha + |u_{i+1} v_{i+1}|^\alpha &= a^\alpha + r^\alpha < (q-1)r^\alpha + r^\alpha \\ &= q r^\alpha = q |u_i v_{i+1}|^\alpha, \text{ for } i = 1, 2, \dots, m-2. \end{aligned}$$

In addition, all links (u_i, v_i) , $i = 2, 3, \dots, m-1$ will be removed as well because

$$\begin{aligned} |v_i v_{i+1}|^\alpha + |v_{i+1} u_i|^\alpha &= a^\alpha + r^\alpha < (q-1)r^\alpha + r^\alpha \\ &= q r^\alpha = q |u_i v_i|^\alpha, \text{ for } i = 1, 2, \dots, m-1. \end{aligned}$$

Therefore after the removal process the shortest path between u_1 and v_1 is $u_1, u_2, \dots, u_{m-1}, v_m, \dots, v_2, v_1$ which renders the stretch factor $O(n)$. It should be noted that this proof does not work if $q = 1$. In (Rajaraman, 2002) it is claimed that removing links using Eq.(4.2.1) in an uncoordinated manner will violate the power spanner property even when $q = 1$. However that is not quite true because when $q = 1$ the only links that are removed are the inefficient ones which cannot be part of the minimum power path between any two nodes and therefore their removal will not harm the power spanner property.

Now we quickly review the work done in finding an upper bound for the power stretch factor of the Yao graph. It is proved that the Yao graph has a length stretch factor $1 + O(\frac{1}{\theta})$ where $\theta = \frac{2\pi}{k}$ and k is the Yao graph's parameter. Thus according to lemma 2 in (Li *et al.*, 2001) the power stretch factor

is at most $(1 + O(\frac{1}{\theta}))^\alpha$. (Bose *et al.*, 2004) show for $0 \leq \theta < \frac{\pi}{4}$ the length stretch factor is bounded by $((\cos(\theta) - \sin(\theta))^{-1})$, hence the power stretch factor is bounded by $((\cos(\theta) - \sin(\theta))^{-\alpha})$. J. Ruppert and R. Seidel proves that $(1 - 2\sin(\frac{\theta}{2}))^{-1}$, where $0 < \theta < \frac{\pi}{3}$, is an upper bound of the length stretch factor, therefore the power stretch factor is bounded by $(1 - 2\sin(\frac{\theta}{2}))^{-\alpha}$. Finally (Li *et al.*, 2001) try to prove that the power stretch factor is bounded by $(1 - (2\sin(\frac{\theta}{2}))^\alpha)^{-1}$, where $0 < \theta < \frac{\pi}{3}$.

We end this section by commenting on the proof of the power stretch factor obtained in (Li *et al.*, 2001). The authors (Li *et al.*, 2001) try to prove that the Yao graph has a power stretch factor bounded by $c_1 = (1 - (2\sin(\frac{\theta}{2}))^\alpha)^{-1}$ for $0 < \theta < \frac{\pi}{3}$. Unfortunately their proof contradicts itself.

They use mathematical induction on the number of links that a path has. It is assumed the following inequality holds as an induction hypothesis

$$P_{YG_k}(\Pi(w, v)) < c_1 |wv|^\alpha, \quad (4.2.2)$$

for any path $\Pi(w, v)$ of l links connecting w and v in the Yao graph, YG_k , and they try to prove that Eq.(4.2.2) holds for any path of $l + 1$ links where $P_{YG_k}(\Pi(w, v))$ is the cost of the path connecting w and v in YG_k . Assume that u and v are not connected in the Yao graph while they are connected in the original topology. Therefore there exists w in the same cone as v such that $|wv| < |uv|$; see Lemma 4.6.⁴ We briefly discuss their Case 2 since it is the place where the error comes in. In this case it is assumed that $\angle vwu < \pi/2$; see Figure 4.18.

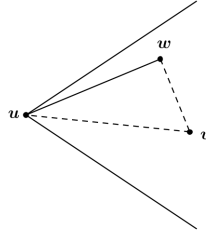


Figure 4.18: The angle $\angle vwu$ is acute.

A path of $l + 1$ links from u to v in G_θ is

$$\Pi(u, v) = (u, w) + \Pi(w, v),$$

where $\Pi(w, v)$ is a path of l links. They wish to show that

$$P_{YG_k}(\Pi(u, v)) < c_1 |uv|^\alpha.$$

⁴Obviously this lemma applies for CBTC cones but it still holds for Yao cones.

From the induction hypothesis we have

$$P_{YG_k}(\Pi(u, v)) \leq |uw|^\alpha + c_1 |wv|^\alpha \quad (4.2.3)$$

The maximum length of $\Pi(w, v)$ occurs when $|uw| = |uv|$ and they use that to write $|wv|$ in terms of $|uv|$ and then substitute in Eq.(4.2.3). However this is erroneous because it contradicts the very assumption that (u, v) is not in the Yao graph. It should be noted that with $\angle vuw < \frac{\pi}{3}$ the mathematical induction argument works as proved in (O' Rourke, 2010), yet the mathematical induction therein is carried over the length of links not over the number of links that a path consists of. Moreover suppose that (u, v) is not in the Yao graph as a result of tie breaking, i.e., initially w and v are at the same distance from u and the algorithm that constructs the Yao graph has chosen only one of them. Since nodes are randomly distributed, this case almost never happens. Therefore we can say that with probability zero the power stretch factor of the Yao graph is c_1 when $\theta < \frac{\pi}{3}$.

Furthermore, c_1 has very weird behaviour since it approaches 1, the optimal value, as the path loss exponent, α , increases. As a matter of fact, when α increases the radio coverage area decreases causing the number of links to decrease and that will definitely increase the stretch factor. In addition, having $c_1 \rightarrow 1$ as the path loss increases implies that $c_1 \rightarrow 1$ as the environment gets complicated which means that the Yao topology performs better in obstructed environments compared with much less obstructed environments! The other power stretch factors obtained for the Yao graph, by contrast, perform worse as the environment gets complicated. Figure 4.19 illustrates this point by comparing the behaviour of c_1 represented by the solid blue line with $c_2 = (1/(1 - 2\sin(\alpha/2)))^\alpha$ represented by the red dashed-dotted line and $c_3 = ((\cos(\theta) - \sin(\theta))^{-\alpha})$ represented by the black dashed line. The green dotted line represents the optimal value and is plotted as a reference.

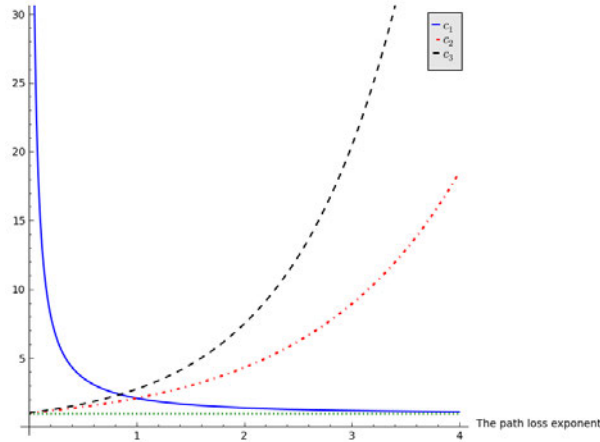


Figure 4.19: While c_1 converges as α increase, c_2 and c_3 diverge. $\theta = \frac{\pi}{6}$.

4.3 Features of XTC and CBTC

In this section we compare the topology produced by XTC, G_{XTC} , with that produced by CBTC, G_θ , in terms of interference, length stretch factor and maximum degree.

Let \bar{G}_θ denote the topology produced by CBTC with angle θ when the asymmetric links are *removed*. We know that the choice of θ determines G_θ and consequently \bar{G}_θ as well. Throughout this section we have chosen θ to be $\pi/2$ and $\pi/3$.

Perhaps the most important feature needed to be well studied is interference. We compare interference of the max-power graph with G_θ and \bar{G}_θ as shown in Figure 4.20. In this figure and all coming ones the labels “CBTC(θ)” and “improved CBTC(θ)” mean G_θ and \bar{G}_θ respectively. Interestingly, while interference of the max-power graph increases linearly with the number of nodes, interference of \bar{G}_θ is *almost* constant. When the asymmetric links are not removed interference grows linearly but with much less slope.

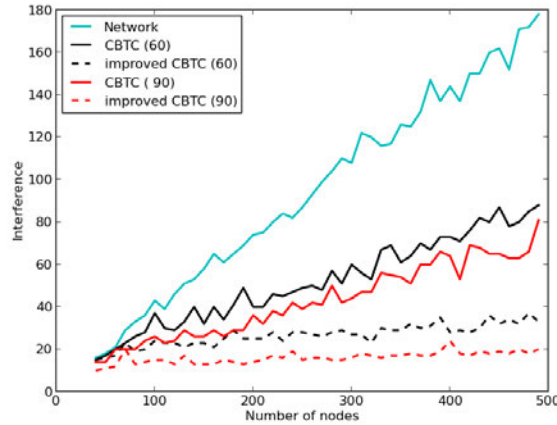


Figure 4.20: Interference is reduced when CBTC is used.

Since \bar{G}_θ outperforms G_θ in terms of interference reduction we have compared \bar{G}_θ with G_{XTC} as Figure 4.21 shows. The interference of the max-power graph is plotted in Figure 4.21(a) as a reference while it is removed in Figure 4.21(b) to allow a closer look to the curves representing interference of the topologies of interest. It is clear that G_{XTC} has the least interference. Moreover it has a constant interference.

A network with less interference is expected to have a high rate of successful exchange of information among its units, yet the amount of information exchanged within the network is not enough to determine the quality of a given wireless *ad hoc* network since there is another imperative factor that has to be considered which is the energy consumption. As we have seen the

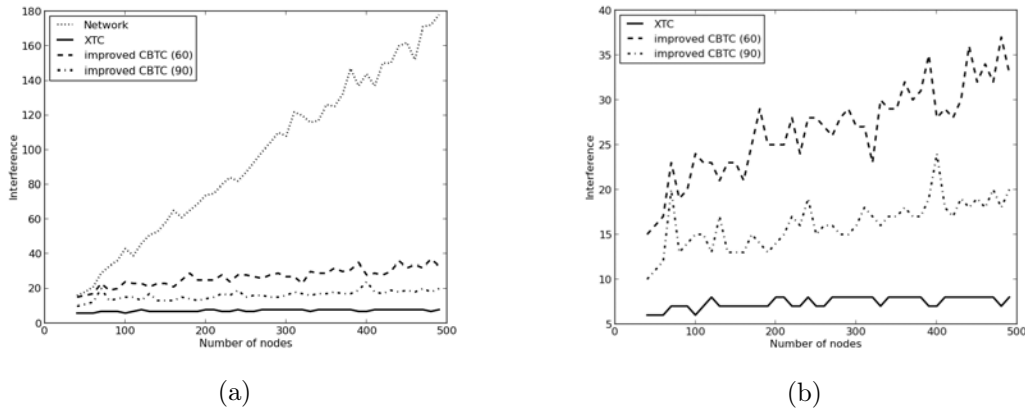


Figure 4.21: Interference of XTC and CBTC.

power stretch factor is a mathematically abstract way to measure energy consumed by the paths in the reduced topology in the worst case compare with the energy consumption of the original topology. From lemma 2 in (Li *et al.*, 2001) we know that if we determine the length stretch factor we can compute the power stretch factor. Thus we have compared the length stretch factor of G_{XTC} with that of \bar{G}_θ as shown in Figure 4.22. From the figure, it is clear that the situation has been reversed here since \bar{G}_θ , especially when $\theta = \pi/3$, outperforms G_{XTC} .

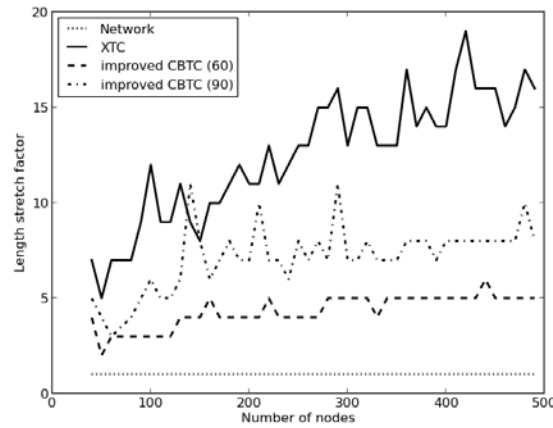


Figure 4.22: Length stretch factor.

Finally we compare the maximum degree in G_{XTC} and \bar{G}_θ as shown in Figure 4.23. Fortunately both topologies have a constant bounded degree. The constant bounded degree property is desired since it decreases node overhead and makes the topology adaptable.

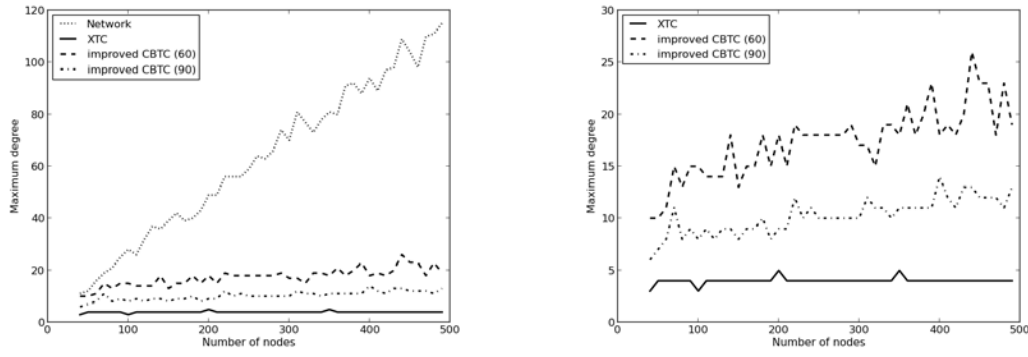


Figure 4.23: Maximum degree.

We conclude the section by stressing that it is not straightforward to determine which one of G_{XTC} and \bar{G}_θ is preferred. An efficient way to do that is through extensive simulation using a network simulator to compare the performance of G_{XTC} with that of \bar{G}_θ with respect to many metrics including the performance of routing protocols, network throughput, network lifetime, resilience to dynamism, etc. One advantage of such a simulation study is to bridge the gap between the practical and theoretical approaches towards wireless *ad hoc* networks. We delay this subject for future research.

Chapter 5

Simulation

This chapter represents the simulation part of the thesis. It contributes to the research area in three respects:

- Firstly, we show that the degree distribution in wireless *ad hoc* networks modelled by the log-normal model is binomial in the presence and absence of the border effect as long as the average degree is not high. In §5.1, we state the problem, show its significance and verify it through extensive simulation.
- Secondly, we show that the node density that guarantees the existence of no isolated node can be used to reliably approximate the node density that ensures connectivity w.h.p. This problem is well stated in §5.2. In addition, its value has been shown and finally verified through simulation.
- Lastly, in §5.3, we show that the log-normal model for wireless *ad hoc* networks is not a reliable model if the environmental parameter, ζ , is at least 6.

5.1 Degree Distribution in Wireless *Ad hoc* Networks

5.1.1 Problem Statement

The distribution of node degree in random graphs is binomial by definition. The degree distribution in random geometric graphs is binomial as well. To show that, let \mathcal{L} be as defined in §2.4.1. Then the probability that \mathcal{L} occurs is $P[\mathcal{L}] = |A_0|/|A|$, where A_0 is the radio coverage area and $A \subset \mathbb{R}^2$, is a 2-dimensional region in which n nodes are distributed uniformly at random. Thus establishing a link is a Bernoulli trial. The degree of a randomly chosen node is, then, a Bernoulli trial repeated $n - 1$ times among them there are k successes where k is the degree of the chosen node. In order for the $n - 1$

Bernoulli trials to be considered as a binomial experiment they have to be independent. In our case the independence of our Bernoulli trials follows from the fact that the event of a node lying within A_0 is independent from the event that any other node lies within or outside A_0 since nodes are identically, independently and uniformly distributed and for a fixed n the degree of a node depends solely on the size of A_0 .

Unfortunately a wireless *ad hoc* networks modelled by the log-normal model is not a pure random graph nor is it a pure random geometric graph. Rather it is mixture of the two. In addition, in reality nodes are distributed in a deployment area with a fixed size and defined boundaries. Assume that the deployment area is $x \times y$, then nodes lying outside the rectangle $(x - 2r_0) \times (y - 2r_0)$, where r_0 is as in §.2.4.1, will have part of their radio coverage area cut by the border of the deployment area. As a result their expected degree is less than that of nodes in the interior region, i.e., inside $(x - 2r_0) \times (y - 2r_0)$. This disturbing phenomenon is known as the *border effect* (Bettstetter, 2004b). Because of the randomness in link formation due to Eq.(2.4.8) and randomness caused by the borders of the deployment area it is not straight forward to decide whether the degree distribution of wireless *ad hoc* networks is binomial. It should be noted that our previous argument for showing that the degree distribution is binomial in random geometric graphs assumes that border effect is absent¹ which is the typical case for random geometric graphs.

Let's state the problem formally. Let $G = (V, E)$ be a graph modelling a wireless *ad hoc* network at a particular instant of time based on the log-normal model, $A \subset \mathbb{R}^2$ be the deployment area and $|A|$ be its size.² Assume that n nodes, each with a transmitting range r_0 , are distributed in A uniformly at random. The degree of a randomly chosen node is, then, a random variable denoted by D . In General $D = D(n, r_0)$, i.e., D depends on the number of nodes (or equivalently density of nodes³) and the transmitting range. For simplicity we suppress the dependence of D on n and r_0 . The question of interest is: *what is the distribution of D ?*

5.1.2 Problem Significance

Knowing the degree distribution is strongly related to the study of connectivity. It is very difficult to obtain the minimum transmitting range that ensures the connectivity of a wireless *ad hoc* network with probability p . If we fix the transmitting range the question becomes what is the minimum node density that ensures connectivity of the network with probability p ?

Let $r_c(p)$ be the minimum transmitting range that ensures connectivity of the network with probability p and $r_i(p)$ be the minimum transmitting range

¹Nodes are assumed to be distributed either in an infinite region (Santi, 2005b, p. 170) or in a *torus* topology as in (Wiki01, 2013).

² A is a measurable set and $|A|$ is a 2-dimensional Lebesgue integral.

³The density, denoted by ρ , is defined as $\rho = n/|A|$ where n is the number of nodes.

that ensures the existence of *no* isolated node in the network with probability p . Let $\rho_c(p)$ be the minimum density that ensures connectivity with probability p and $\rho_i(p)$ be the minimum density that ensures the existence of *no* isolated node with probability p . Obtaining $r_c(p)$ or $\rho_c(p)$ analytically is very complicated while obtaining $r_i(p)$ or $\rho_i(p)$ is relatively easier so $r_i(p)$ respectively $\rho_i(p)$ is obtained instead and used as an approximation to $r_c(p)$ respectively $\rho_c(p)$ (Bettstetter, 2004b; Bettstetter and Hartmann, 2005). The reliability of this approximation is the subject of §.5.2.

Suppose that the probability distribution of D is known. Assume that F_D is the probability mass function of D . Then the probability that a randomly chosen node is isolated is $P[D = 0] = F_D(0)$. Assume the statistical independence of events of isolated nodes⁴ then the probability that there is no isolated node in the network is given by

$$P[D > 0] = (1 - P[D = 0])^n = \hat{p},$$

where $0 \leq \hat{p} \leq 1$. In general $P[D > 0]$ is a function of r_0 and ρ and it is solved for either r_0 or respectively ρ for a given \hat{p} to obtain an analytical expression for $r_i(\hat{p})$ or respectively $\rho_i(\hat{p})$. Once $r_i(\hat{p})$ or respectively $\rho_i(\hat{p})$ is obtained, it can be used to approximate $r_c(\hat{p})$ or respectively $\rho_c(\hat{p})$ under certain conditions (Bettstetter, 2004b; Bettstetter and Hartmann, 2005).

The essence of the aforementioned argument is that determining the degree distribution is closely related to the study of connectivity in wireless *ad hoc* networks. Needless to say that connectivity is, perhaps, the most important characteristic that the reduced topology has to satisfy.

5.1.3 Tests of Hypotheses

To verify whether the degree distribution in wireless *ad hoc* networks is binomial, we use tests of hypotheses. An introduction to test of hypotheses can be found in (Ott and Longnecker, 2008; Crawshaw and Chambers, 2001). Here, we briefly recall the general concepts.

Our problem can be stated in terms of tests of hypotheses as follows:

H_0 : The degree distribution of wireless *ad hoc* networks in the presence/absence of border effect is binomial.

H_a : The degree distribution is not binomial.

We refer to H_0 as the *null hypotheses* and H_a as the *alternative hypotheses*. The kind of test we are doing here is referred to as *two-sided* test or *two-tailed* test. After performing the test we either accept or reject H_0 in favour of H_a which means that the choice of H_a influences our decision with regards to H_0 .

Hypotheses tests are generally of two types: *parametric* tests and *non parametric* or more precisely *distribution free* tests. Parametric tests make certain

⁴The events of isolation are almost independent from node to node if $n \gg 1$ and $A_0/A \ll 1$ where A_0 is the radio coverage area.

assumptions about data while non parametric tests do not make any assumptions about data; hence the name *distribution free*. In general parametric tests are more *powerful* if their assumptions are met. By “power” of a statistical test we mean its ability to detect the difference between the empirical (observed) probability distribution and the expected (theoretical) probability distribution when the difference actually exists (National Institute of Standards & Technology, 2008). However they can be very misleading if their assumptions are violated. Non parametric tests are more *robust*. By “robustness” of a statistical test we mean that the test performs well under a wide range of distributional assumptions (National Institute of Standards & Technology, 2008).

We use the chi-square test (a representative of the parametric family) and the Kolmogorov-Smirnov test (a representative of the non parametric family). We write *ks-test* as shorthand for Kolmogorov-Smirnov test.

The chi-square statistic is calculated as follows (Spiegel *et al.*, 2000):

$$\chi = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}, \quad (5.1.1)$$

where O_i is the observed frequency of bin (cell) i and E_i is its expected frequency.

The ks-test statistic is defined as follows (Wiki02, 2013):

$$K = \sup_x |F_k(x) - F(x)|, \quad (5.1.2)$$

where $F(\cdot)$ is the expected distribution function and $F_k(\cdot)$ is the empirical distribution function calculated as follows:

$$F_k(x) = \frac{1}{k} \sum_{i=1}^k I(X_i \leq x),$$

where

$$I(X_i \leq x) = \begin{cases} 1 & \text{if } X_i \leq x \\ 0 & \text{otherwise.} \end{cases}$$

As Eq.(5.1.1) suggests, the chi-square test detects the difference between the empirical and the expected distribution by calculating the total relative deviation of the observations from the expected values while as Eq.(5.1.2) suggests, the ks-test detects the difference by calculating the maximum difference between the empirical and the expected distribution.

If H_0 is rejected while it is true, we say *type I error* has been made. On the other hand if H_0 is accepted while it should be rejected we say *type II error* has been made (Spiegel *et al.*, 2000). The maximum probability at which we are willing to risk a type I error is called the significance level. We use 0.05 significance level. Attempts towards minimizing one type of error is associated

with an increase in the other type of error. In practice one type of error is more serious so it is minimized. In our study a type II error is more serious since it will result in rendering all theoretical studies that assume H_0 irrelevant and erroneous.

5.1.4 The Simulation

We distribute n nodes uniformly at random in a square area A of side l . We assign to each node a transmitting range r_0 . We establish links between all possible pair of nodes according to Eq.(2.4.8).

It is worth noting that links formed according to Eq.(2.4.8) are not symmetric in general. In practice links are symmetric. Therefore we make all links formed according to Eq.(2.4.8) symmetric by adding (v, u) to E in case $(u, v) \in E$ and vice versa. After constructing the network topology, we calculate the empirical probabilities as follows:

$$P_i[D = k] = \frac{n_k}{n},$$

for $k = 0, 1, \dots, n-1$ and $n_k = |\{u \in V \mid |N_u| = k\}|$ where N_u is the set of neighbours of u . Simply, n_k is the number of nodes with degree k . The subscript i denotes the simulation run i . We repeat the process 1000 times.

The expected probabilities are generated from the probability mass function

$$P[D = k] = \binom{n-1}{k} p^k (1-p)^{n-k-1}, \quad (5.1.3)$$

where $p = P[\mathcal{L}]$.

To generate the expected probabilities, we need to calculate p as follows. Since the n nodes are distributed uniformly at random, the distance between two randomly chosen nodes is a random variable. Let R denote such a random variable. The probability density function of R can be approximated by (Stuedi *et al.*, 2005):

$$F_R(r) \approx \frac{2r^3 - 8lr^2 + 2\pi l^2 r}{l^4}, \quad (5.1.4)$$

where $0 \leq r \leq l$.

Therefore the expected value of the probability of a link event, $P[\mathcal{L}]$, is given by:

$$P[\mathcal{L}] = \int_{r=0}^l F_R(r) P[\mathcal{L}|r] dr = p, \quad (5.1.5)$$

with $P[\mathcal{L}|r]$ as in Eq.(2.4.8).

Although Eq.(5.1.5) can be integrated numerically to obtain an approximation to p , we calculate p empirically as follows: we construct network topologies $G_i = (V_i, E_i)$ where $i = 0, 1, \dots, N$ denotes the simulation run and

$ A $	ζ	r_0	$P[\mathcal{L}]$ (theoretical)	$P[\mathcal{L}]$ (empirical)
10^2	1.0	1.0	0.0316	0.0315
10^2	1.0	2.0	0.1138	0.1137
50^2	2.0	10.0	0.1410	0.1423
50^2	3.0	5.0	0.0611	0.0605
100^2	1.0	5.0	0.0083	0.0084
100^2	2.0	10.0	0.0414	0.0411

Table 5.1: Values for $P[\mathcal{L}]$ calculated by solving Eq.(5.1.5) numerically and by using Eq.(5.1.6) with Eq.(5.1.7) (empirical probability) for different area sizes, different transmitting ranges and different environments.

N is a large integer. Then the probability of a link event in the simulation run i is calculated as follows:

$$P_i[\mathcal{L}] = p_i = \frac{|E_i|}{\binom{n}{2}}, \quad (5.1.6)$$

where $\binom{n}{2}$ is the number of all possible links. Then according to *law of large numbers* (Spiegel *et al.*, 2000) with probability one the following holds:

$$\frac{1}{N} \sum_{i=1}^N p_i \longrightarrow p \quad \text{as} \quad N \longrightarrow \infty. \quad (5.1.7)$$

Eq.(5.1.7) states that if N is sufficiently large, the empirical probability approaches the theoretical probability. We have found that $N = 100$ is large enough to generate p accurately. We integrate Eq.(5.1.5) numerically and compare values of $P[\mathcal{L}]$ obtained from Eq.(5.1.5) with values of $P[\mathcal{L}]$ obtained from Eq.(5.1.7) (simulation) and we find that they are almost equal. Table 5.1 presents the result for different area sizes, different environmental parameters and different transmitting ranges.

After calculating the expected and empirical values we calculate the chi-square statistic, χ , according to Eq.(5.1.1) and the ks-statistic, K , according to Eq.(5.1.2). We reject H_0 if the p value of the test is less than 0.05. The p value is the probability of getting a test statistic as extreme as the one we have got if H_0 is true (Spiegel *et al.*, 2000). If it is very small (less than 0.05 in our case) then it is unlikely to get a sample like the one we have if H_0 is true and hence we reject it.

An alternative approach to ours for verifying H_0 is to generate the observed and the expected values as we explained and then do the test. Then generate another set of the observed and expected values and perform the test once again and so on. Assume that the process is repeated N times, out of which

H_0 is accepted k times. Then k/N defines a “level” of certainty on accepting H_0 at the 0.05 significance level.

We choose the first approach because the second one generates the observed values only one at a time while the first approach generates such a huge sample that the difference is easily detected when it actually exists. Our sample is based on $1000n$ observations which is a quite large number since typically we choose $n = 50, 100, 200$.

As mentioned earlier the log-normal model is neither equivalent to the random graph nor is it equivalent to random geometric graph. Therefore even in the absence of the border effect the degree distribution is not necessarily binomial. To be able to simulate the case when the border effect is negligible, we consider an observation area B inside the simulation area A where the distance between the border of B to the border of A is at least r_ϵ where r_ϵ satisfies

$$P[\mathcal{L}|r_\epsilon] \leq \epsilon \quad (5.1.8)$$

or simply r_ϵ is the distance at which the probability of a link event is at most ϵ . If $\epsilon \ll 1$, a negligible portion of the radio coverage area of nodes lying on the boundary region of B will be outside A . Obviously *only* nodes inside B are considered for statistics, yet they are allowed to establish links with any node inside A . We choose $\epsilon = 0.05$ and solve Eq.(5.1.8) numerically to obtain r_ϵ .

5.1.5 Statistical Considerations

One major disadvantage of the ks-test is that if one of the parameters of the expected probability distribution is estimated from the sample, the test is no longer valid (National Institute of Standards & Technology, 2008). To avoid this problem, we estimate $P[\mathcal{L}]$ from a different data set.

The test is designed to only test *continuous* distributions, hence it cannot be used to test a binomial distribution. However if n is large and p is not very small nor is it very large, the binomial distribution $B(n, p)$ can be approximated by the normal distribution $N(np, \sqrt{np(1-p)})$. The approximation is good if $np > 5$ and $n(1-p) > 5$ (Ott and Longnecker, 2008) since if $np < 5$, which means that $p \ll 1$ and n is modest, the binomial distribution is seriously skewed left; while if $n(1-p) < 5$, which means that $p \approx 1$ and n is modest, the binomial distribution is seriously skewed right. In both cases the symmetry of the normal distribution is violated and hence the approximation is no longer good. Figure 5.1 illustrates this point.

For wireless *ad hoc* networks, especially wireless sensor networks, n is fairly large and $P[\mathcal{L}]$ is neither very small nor very large. Since $P[\mathcal{L}] \ll 1$ results in a disconnected network consisting of many small connected sub-networks. On the other hand, $P[\mathcal{L}] \approx 1$ results in high interference and high energy consumption. Therefore it is safe to assume the normal approximation to the

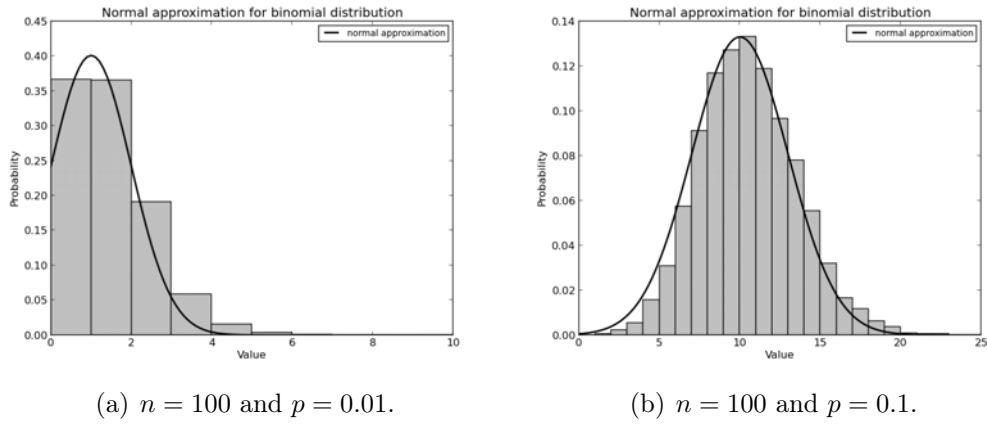


Figure 5.1: Bad and good normal approximations to the binomial distributions.

binomial distribution is good and hence we can use the ks-test to test degree distribution.

Although the chi-square test is the most popular, it can be very misleading and inaccurate if one of its assumptions and conditions is violated. The statistic χ in Eq.(5.1.1) follows the chi-square distribution if the sample size, n , is large. The sample size should be chosen such that 80% of bins satisfy $np_i > 5$ where p_i is the expected value of bin i (Ott and Longnecker, 2008). Unfortunately any attempt at deciding the optimal n is hopeless (Ott and Longnecker, 2008). To ensure this condition we combine bins with frequencies less than 5. However if our observed values are very small the combination process can render the sample size so small that χ is no longer distributed according to the chi-square distribution. Unfortunately despite ensuring the chi-square conditions have been met, it tends to always reject H_0 . On the contrary, the ks-test gives reasonable results.

A common assumption is the statistical independence of the observations. If the observations are correlated neither of the two tests is valid. The assumption of the independence of the observations is reasonable and can be justified since nodes are independent identically distributed.

5.1.6 Results and Discussions

We conduct the simulation for different values of n , different values of $|A|$ and different values of ζ . Table 5.2 reports the result of our simulation when $|A| = 100$ and $n = 50, 100$ and 200 . The environmental parameter ζ takes different values as shown in the table. We set $r_0 = 0.2$ at the beginning, then we increase r_0 by 0.2 at a time. We stop when $|l/\sqrt{\pi} - r_0| < 0.2$ where $l/\sqrt{\pi}$ is the maximum value that r_0 can take and l is the side of the simulation area. Our conclusions based on the simulation are listed below.

1. The degree distribution in wireless *ad hoc* networks in the presence of

Density	ζ	r_0	A_0/A	$E[D]$	$E[D]/n$
$\rho = 0.5$	0.5	3.8	0.454	15.8 ± 1.3	0.316 ± 0.026
	1	3.6	0.407	15.1 ± 1.3	0.302 ± 0.026
	2	4.4	0.608	22.0 ± 1.5	0.44 ± 0.03
	3	5.6	0.985	28.3 ± 1.2	0.566 ± 0.024
$\rho = 1$	1	2.8	0.246	20.2 ± 1.2	0.202 ± 0.012
	2	2.6	0.212	25.7 ± 1.2	0.257 ± 0.012
	3	2.6	0.212	25.7 ± 1.2	0.257 ± 0.012
	4	3.0	0.283	34.2 ± 1.2	0.342 ± 0.012
	5	3.8	0.454	43.7 ± 1.2	0.437 ± 0.012
	6	5.6	0.985	54.8 ± 1.1	0.548 ± 0.011
$\rho = 2$	1	2.2	0.152	26.8 ± 1.0	0.134 ± 0.005
	2	2.0	0.126	28.1 ± 1.1	0.141 ± 0.006
	3	1.6	0.080	26.1 ± 0.940	0.131 ± 0.005
	4	1.4	0.060	28.8 ± 0.920	0.143 ± 0.005
	5	1.4	0.060	36.6 ± 0.970	0.183 ± 0.005
	6	2.0	0.126	58.3 ± 1.2	0.292 ± 0.006
	10	5.6	0.985	106.7 ± 1.0	0.534 ± 0.005

Table 5.2: Simulation results when $|A| = 100$, $n = 50, 100, 200$ and ζ is as shown. The table shows the result for the minimum transmitting range at which H_0 is rejected.

border effect is binomial if the average degree is not “large”. The question: *how large is “large”* depends on *i)* the environmental parameter *ii)* the radio coverage area and *iii)* the node density. Therefore there is no fixed threshold such that the degree distribution is binomial if the average degree does not exceed that threshold.

The small average degree is reflected by small radio coverage area and low node density. We think that it is valid to assume the degree distribution is binomial when the connectivity of wireless *ad hoc* networks is analysed since the network becomes connected at low average degrees. Furthermore in practice the network devices are unlikely to have large radio coverage area. We conclude that if $|A_0|/|A| \ll 1$ or $\rho \ll 1$ the border effect is negligible. For example when $|A| = 100$, $\rho = 0.2$ and $\zeta = 1$, A_0 can be as large as the whole simulation area, A , without distorting the binomial distribution of degree. In addition, we find that if $|A_0|/|A| \approx 0$, the degree distribution is binomial no matter how large ρ is. Figure 5.2 shows examples of the empirical probability mass function and the expected probability mass function when H_0 is not rejected.

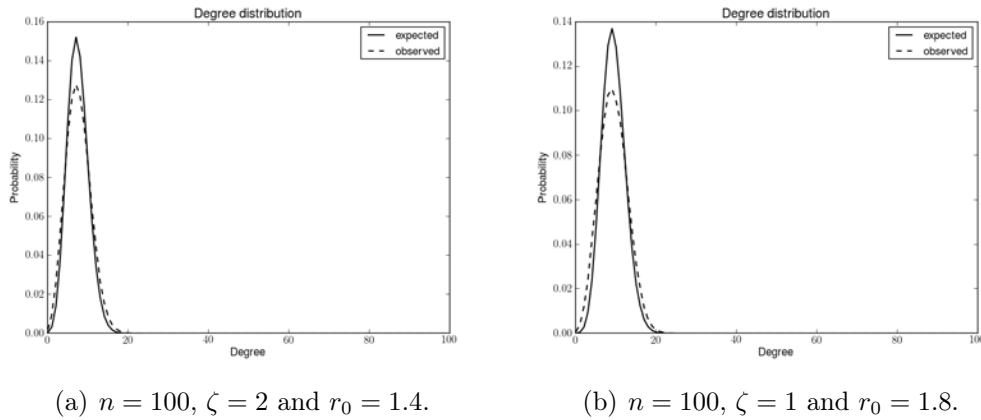


Figure 5.2: Empirical and expected probability mass functions when H_0 is not rejected.

- For a fixed ζ , as ρ increases, $|A_0|$ has to decrease in order for H_0 to be accepted. This is quite obvious since when ρ increases, more nodes are affected by the distorting effect of the border. To offset this effect $|A_0|$ decreases. Consequently the ratio of the average degree to n decreases despite the fact that the density, ρ , has increased. For example when $\zeta = 1$; the ratio is 0.302 when $\rho = 0.5$, it is 0.202 when $\rho = 1$ and it is 0.134 when $\rho = 2$; see Table 5.2.

3. In the presence of border effect, for a fixed ρ and $\zeta = 0$,⁵ we know that there is a transmitting range, refer to it as \tilde{r}_0 , such that if $r_0 = \tilde{r}_0$, H_0 is accepted and it is rejected when $r_0 = \tilde{r}_0 + \epsilon$ for any $\epsilon > 0$. In other words, \tilde{r}_0 is the maximum transmitting range at which the border effect does not alter the degree distribution of the random geometric graph. As ζ increases, it distorts the regularity of the random geometric graph by removing some short links and adding some long ones. As a result r_0 decreases to balance the distorting effect of the increasing ζ . On the other hand, as ζ increases, the correlation between links decreases, see §5.3, and the graph starts to become similar to a random graph and hence the border effect becomes more tolerable. Therefore there must exist $\tilde{\zeta}$ such that if $\zeta < \tilde{\zeta}$, r_0 decreases as ζ increases in order for H_0 to be accepted and if $\zeta > \tilde{\zeta}$, r_0 starts to increase as ζ increases without resulting in rejection of H_0 . When ζ is very large the graph is so similar to a random graph that the border effect does not matter any more. As a result A_0 can be as large as the whole simulation area. For example as Table 5.2 shows when $\rho = 1$ and $\zeta = 1$, we do not reject H_0 unless $r_0 \geq 2.8$. As ζ increase r_0 decreases. However there is $3 < \tilde{\zeta} < 4$ such that if $\zeta > \tilde{\zeta}$ as ζ increases, the transmitting range, r_0 , can increase as well without resulting in the rejection of H_0 . In other words, when $\zeta > \tilde{\zeta}$ the log-normal graph becomes *more* random graph than random geometric graph and thus the border effect becomes of lesser significance. There is $\hat{\zeta}$ such that if $\zeta > \hat{\zeta}$ the border effect is no longer a problem and thus A_0 can be as large as A . Figure 5.4 shows the typical plot of ζ and the maximum transmitting range r_0 at which H_0 is not rejected for a given ρ . When $\zeta > \hat{\zeta}$ the transmitting range can take its maximum values, $R = l/\sqrt{\pi}$, without resulting in the rejection of H_0 . It should be noted that when $\rho = 2$ and $\zeta = 4$, the minimum transmitting range at which we reject H_0 is $r_0 = 1.4$ which is the same as the transmitting range when $\zeta = 5$, yet the average degree when $\zeta = 5$ is 36.6 while it is 28.8 when $\zeta = 4$ which implies that as ζ increases connectivity improves. When ζ increases some short links are removed but some long links are added. Improvement of connectivity as ζ increases implies that the number of long links added to the network is larger than the number of short links removed. Figure 5.3 shows a network consisting of 100 nodes with $r_0 = 1$. As the figure shows, when ζ increases connectivity improves which validates our conclusion based on our simulation.
4. In the absence of the border effect, the randomness in link formation caused by the log-normal model does not alter the binomial distribution for node degree. This is a valuable result for theoretical studies as they usually ignore the border effect. Our simulation results allow them to

⁵Remember that when $\zeta = 0$ the log-normal model is equivalent to the random geometric graph model.

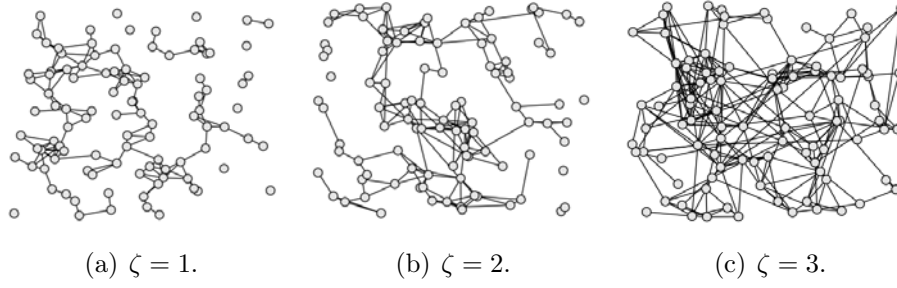


Figure 5.3: Connectivity improves as ζ increases.

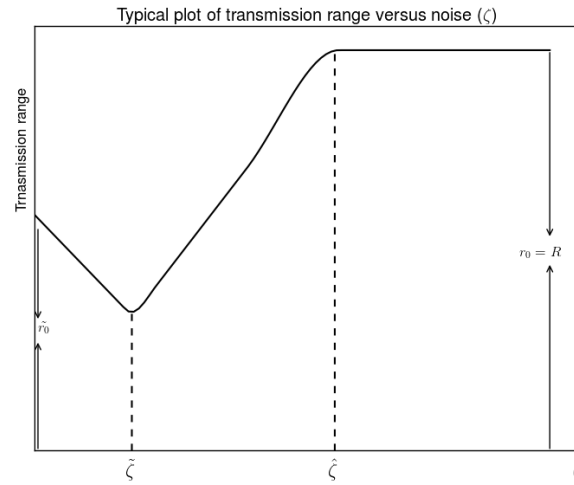
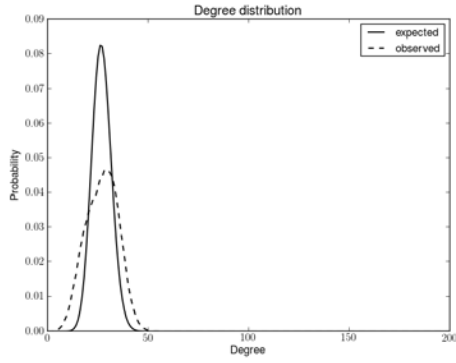


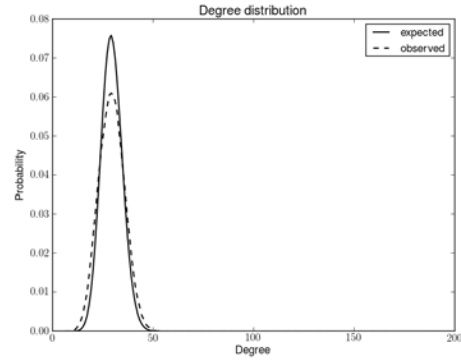
Figure 5.4: Typical plot of ζ and the maximum transmitting range at which H_0 is not rejected. R is the maximum of r_0 .

safely assume that the degree distribution is binomial. Figure 5.5 and Figure 5.6 show the the empirical probability mass function and the expected one at the point when we reject H_0 for the first time in the presence of border effect and when the border effect is avoided. We note that when the border effect is avoided the difference between the two probability functions is so small that it does not result in rejecting H_0 .

5. The increase in r_0 distorts the degree distribution seriously when the border effect is present as Figure 5.7 shows. When $r_0 = 1$ the empirical probability mass function and the expected probability mass function are almost identical, when $r_0 = 2$ they deviate from each other and when $r_0 = 3$ the deviation is very prominent, yet it does not result in rejecting H_0 .

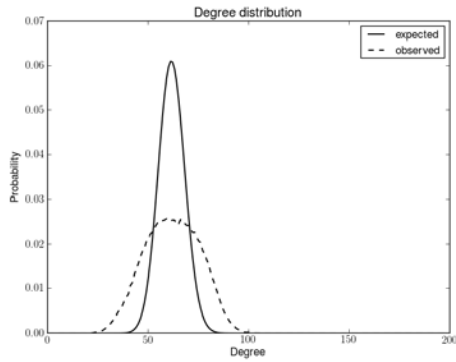


(a) Border effect is present

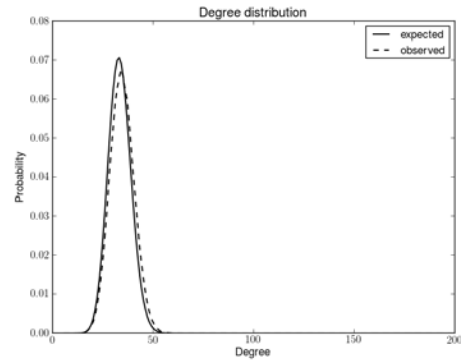


(b) Border effect is avoided

Figure 5.5: $|A| = 100$, $\rho = 2$, $\zeta = 1$ and $r_0 = 2.2$.

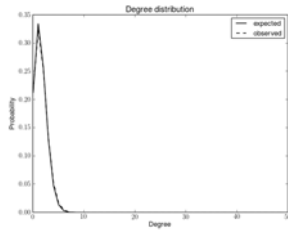


(a) Border effect is present

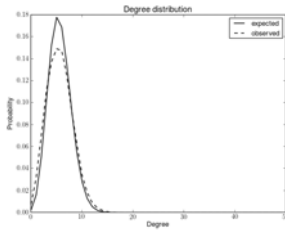


(b) Border effect is avoided

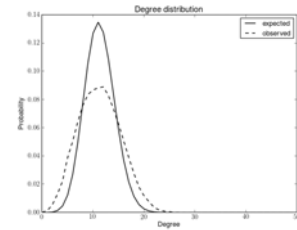
Figure 5.6: $|A| = 100$, $\rho = 2$, $\zeta = 3$ and $r_0 = 1.6$.



(a) $r_0 = 1.0$.



(b) $r_0 = 2.0$.



(c) $r_0 = 3.0$

Figure 5.7: The increase in r_0 distorts the degree distribution seriously.

5.2 Connectivity and Minimum Node Degree

5.2.1 Statement of the Problem

In this section we show that connectivity of wireless *ad hoc* networks can be guaranteed w.h.p if there is no isolated node in the network. As we mentioned, $\rho_i(p)$ is calculated and used as an approximation to $\rho_c(p)$. We know that $\rho_i(p)$ is a lower bound for $\rho_c(p)$. The subject of this section is to investigate the reliability of this approximation and the tightness of this lower bound. In short, this section answers the following questions:

i) Is the lower bound *tight*? *ii)* If it is so, is it *always* tight or it is *only* tight under certain conditions?

Without further proceeding, we present a theorem that is related to the problem of our interest. Penrose (Penrose, 1999) proves the following theorem:

Theorem 5.1. *For a random geometric graph with n nodes distributed uniformly at random, let $r_c^k(n)$, receptively $r_i^k(n)$ be the minimum transmitting range at which the graph is k -connected, respectively has a minimum node degree k . Then*

$$P[r_c^k(n) = r_i^k(n)] \rightarrow 1 \text{ as } n \rightarrow \infty. \quad (5.2.1)$$

This theorem is of theoretical value because it guarantees *k-connectivity* by ensuring that the minimum node degree is k provided that the number of nodes or equivalently the density is sufficiently large. $r_i^k(n)$ can be calculated analytically given that the degree distribution is known as done in (Bettstetter, 2004a,b, 2002). The degree distribution of wireless *ad hoc* network is discussed in §.5.1.

In the case when $k = 1$, we can ensure connectivity w.h.p by ensuring that there is no isolated node. The intuition behind this point is explained in §.5.1.2.

Theorem 5.1 can be interpreted as follows: suppose that a wireless *ad hoc* network is modelled by a random geometric graph. In addition, assume that all nodes have the same transmission capabilities. We increase the transmitting range synchronously at each node. The theorem states that: if n is sufficiently large, then with high probability the network becomes connected *at the same moment* when the transmitting range is large enough to ensure that there is no isolated node.

In terms of graph theory, Theorem 5.1 has the following interpretation: in random geometric graphs, if n is sufficiently large, the link that is large enough to connect the last isolated node to its nearest neighbour and the longest link in the Euclidean minimum spanning tree are *equal*. The link that connects the last isolated node to its nearest neighbour can be seen as the minimum transmitting range that ensures the absence of isolated nodes and the longest

link in the Euclidean minimum spanning tree can be seen as the minimum transmitting range that ensures connectivity.

Unfortunately Theorem 5.1 is not of great value if wireless *ad hoc* networks are modelled using the log-normal model since the log-normal model is not equivalent to the random geometric graph model. Furthermore, the theorem fixes n and increases r_0 at each node synchronously. However we are actually more interested in the reverse: fix r_0 and increase n because it is more practical. In practice, it is technically infeasible to increase r_0 as one wishes. By contrast, increasing n is feasible very often: increasing n means increasing the number of network devices operating in a given area.

5.2.2 The Simulation

We answer the two questions of §5.2.1 through simulation. We consider a simulation area S and observation area A where the distance from the border of A to the border of S is at least r_c where r_c is as in Eq.(5.1.8). On this setting the border effect is negligible. The number of nodes in S , denoted by M , follows the Poisson distribution with intensity $\lambda = \rho |S|$. We distribute the M nodes in S uniformly at random. We then form links between all possible pairs of nodes according to Eq.(2.4.8). Then for a given ρ we check whether this topology: *i*) is connected *ii*) has no isolated node.

We repeat the process N times and calculate the percentage of connected topologies and the percentage of topologies with no isolated nodes. According to law of large numbers these percentages converge to the actual probabilities with probability 1 if $N \rightarrow \infty$. We increase ρ and repeat the whole process. We stop when the percentage of topologies with no isolated nodes is within the interval 0.99 ± 0.005 .

It is obvious that we do not consider nodes located outside A for statistics, yet we can use them as relays to connect nodes in A . We check the connectivity of nodes in A as follows:

1. There are many ways to check connectivity. One way is a simple flooding algorithm as the one used in (Glauche *et al.*, 2003). This algorithm chooses a node at random, say u , tags it and associates with it a set which represents the *component* that u belongs to. Then it appends the neighbours of u to the component and tags them. For each newly added neighbour the algorithm appends its neighbours to the component and tags them and so on until there are no more nodes to append to the component.⁶ If the order of the component is the same as the order of the graph, then the graph is connected otherwise it is not. It should be noted that this simple flooding approach cannot be used to check the

⁶The original version of the algorithm is used to determine the *giant component*. In this case we continue running the algorithm until there is no untagged node left. To determine connectivity it is obvious that we do not have to run the algorithm for all untagged nodes.

connectivity of nodes in A because there might exist a node in A which does not belong to the component of another node in A , yet they can still be connected via a path containing at least one node which is in S but not in A . Another way for checking connectivity is to check whether there is a path between any possible pair of nodes. Alternatively, one can check whether there is a path from a single node to all other nodes. Note that the last two ways can be used in our case. We have chosen the second one because it is less computationally expensive.⁷

2. To check whether a path from a source node to a target node exists we implement Dijkstra's algorithm (Cormen *et al.*, 2001; Wiki03, 2013). Dijkstra's algorithm finds a shortest path from a source node to all other nodes, i.e., it constructs a tree rooted at the source node and spanning all other nodes where the distance from the source node to any other node is the minimum. It is obvious that the algorithm can be used to find a shortest path from a single node to a single target.

We consider a node in A as an isolated node if there is no path in S connecting it to another node in A , i.e., if it cannot communicate with other nodes in A . Note that according to this *definition* of isolation, a node in A is considered isolated even though it has a neighbour in S as long as there is no path connecting it to another node in A .

5.2.3 Results and Discussions

The main result of our simulation is the following:

$$\lim_{p \rightarrow 1} |\rho_c(p) - \rho_i(p)| \rightarrow 0. \quad (5.2.2)$$

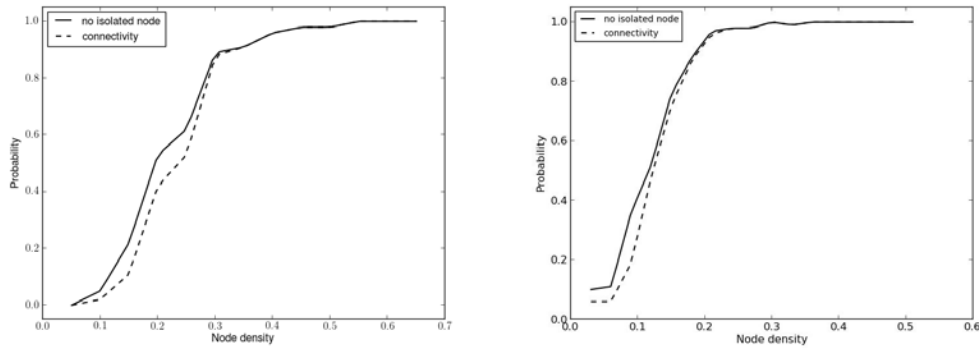
Fortunately, since we require the network to be connected w.h.p. This result is very useful. Under this condition, i.e., $p \approx 1$, we can use $\rho_i(p)$ as a reliable approximation to $\rho_c(p)$. This result answers the two questions that we have raised in §5.2.1.

The result has been confirmed by different simulation runs for different area sizes, different transmitting ranges, r_0 , and different environments, ζ . Figure 5.8 reports the result of the simulation for two different settings as shown in the caption.

Another result of our simulation is verifying one analytical expression obtained in (Bettstetter and Hartmann, 2005) for $\rho_i(p)$. The expression is as follows:

$$\rho_i(p) = -\frac{1}{2\pi\Upsilon} W_{-1} \frac{2\pi\Upsilon \ln(p)}{|A|}, \quad (5.2.3)$$

⁷Note that the computational cost concerns us because we run the code for very large number of times to reliably generate the empirical probabilities. In addition, we increment ρ by small value at each simulation run and the number of nodes in A can be very large depending on ρ and $|S|$.



(a) $|A| = 100$, $n = 200$, $\zeta = 2$ and $r_0 = 2.0$. (b) $|A| = 100$, $n = 200$, $\zeta = 3$ and $r_0 = 2.0$.

Figure 5.8: The probability that the network is connected converges to the probability that the network has no isolated node when $P[D > 0] \approx 1$.

where $W_{-1}(\cdot)$ denotes the real-valued, non-principal branch of the LambertW function, $|A|$ is the size of the observation area and Υ is given by:

$$\Upsilon = \int_0^{\infty} P[\mathcal{L}|r] r dr. \quad (5.2.4)$$

There is overall agreement between the analytical values obtained for $\rho_i(p)$ using Eq.(5.2.3) and values obtained through simulation. This agreement confirms the analytical expression. Table 5.3 presents the analytical and simulated values for $\rho_i(p)$ when $p = 0.99 \pm 0.005$. It should be noted that the simulated values presented in the table are not unique since there is randomness involved when calculating them: *i*) randomness in M which follows the Poisson distribution, *ii*) randomness in node positions, and *iii*) randomness in link formation. Therefore it is possible to obtain better or worse values than those presented in Table 5.3. However, based on many simulation runs we think that the simulated values of $\rho_i(p)$ are normally distributed with mean equal to the value obtained by Eq.(5.2.3) and standard deviation σ which is very small.

From Table 5.3 it is clear that as ζ increases, $\rho_i(p)$ decreases if r_0 is fixed. This observation confirms the result obtained in §5.1.6 that an increase in ζ improves connectivity.

Figure 5.9 shows a plot of the transmitting range and the node density that is required to ensure the absence of isolated nodes with probability 0.99.

A	ζ	r_0	ρ_i (analytical)	ρ_i (simulated)
10^2	1	2.5	0.377	0.370
10^2	2	2.5	0.262	0.245
10^2	3	2.5	0.142	0.125
10^3	1	5	0.106	0.095
10^3	2	5	0.074	0.074
10^3	3	5	0.040	0.040
10^3	3	2.5	0.193	0.190
10^3	4	2.5	0.084	0.081
10^4	2	10	0.020	0.019
10^4	3	10	0.0114	0.0115

Table 5.3: Analytical and simulated values obtained for the density required to ensure the existence of no isolated node in the network. $p = 0.99 \pm 0.005$.

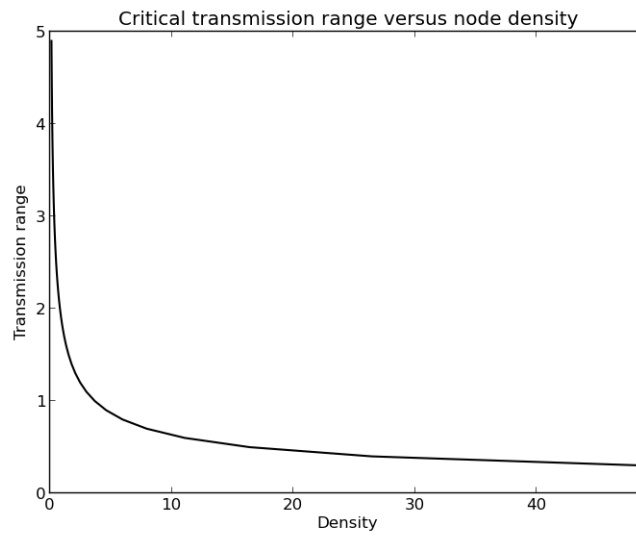


Figure 5.9: Plot of r_0 versus $\rho_i(p)$ where $p = 0.99$, $|A| = 10^3$ and $\zeta = 1$.

5.3 The Range of the Environmental Parameter

The environmental parameter, ζ , plays a key role in the log-normal model since it captures the environment in which a wireless *ad hoc* network is operating. We know that $\zeta = \sigma/\alpha$. Values of σ can vary from 0 up to 12 in a heavily-built environment and values for α can vary from 2 up to 6 (Hekmat, 2006a). Therefore ζ can *theoretically* vary from 0 up to 6, however it is very unlikely that ζ takes values larger than 3 (Hekmat, 2006a) because large values of ζ correspond to large values of σ and small values of α . However large values of σ are due to very irregular environments which, in turn, results in large values of α . Thus in reality σ and α cannot be decoupled.

It is worth mentioning that measurements used to determine the aforementioned ranges for σ and α are done in frequencies lower than typical frequencies used in wireless *ad hoc* networks (Hekmat, 2006a). Consequently ranges of σ and α could be different if wireless *ad hoc* networks are considered.

Partly because there are no reliable measurements⁸ available (Hekmat, 2006a) to determine the typical values of σ and α for typical environments for wireless *ad hoc* networks and partly because *theoretical* values of ζ cannot be larger than 6, we investigate the log-normal graph when $\zeta \geq 6$.

We know that, depending on ζ , the log-normal graph is somewhere between the random geometric graph and the random graph. Our main conclusion for this section is the following:

When $\zeta \geq 6$, a log-normal graph becomes so similar to a random graph that it is no longer a reliable model for wireless ad hoc networks.

In random geometric graphs, link probability is a simple step function given by Eq.(2.4.6). In random graphs link probability is just a horizontal line, i.e., it does not depend on distance. From the plot of link probability given by Eq.(2.4.8) versus the normalised distance, it is obvious that as ζ increases the curve straightens and starts to look similar to a horizontal line meaning that the dependence on distance reduces. When $\zeta = 6$, we can see that the link probability depends *lightly* on the normalised distance; see Figure 5.10.

As a result of lack of dependence on distance the border effect becomes less significant and it keeps losing significance up to a point where it is no longer an issue unless $\rho \rightarrow \infty$. Thus if $\zeta > 6$ we expect that $|A_0|/|A|$ can be as large as 1 without distorting the degree distribution unless $\rho \rightarrow \infty$. This result may seem to contradict the result obtained in Table 5.2 when $\zeta = 6$ and $\rho = 2$ the ratio $|A_0|/|A|$ is 0.126. However the density required to achieve the event that there is no isolated node with probability 0.99 for a network with these characteristics is only 0.00756 which is approximately 264 times smaller than

⁸In (Hekmat, 2006a) there are measurements to determine σ and α for wireless *ad hoc* networks. Unfortunately according to (Hekmat, 2006a) their measurements are not extensive and not very reliable and thus cannot be used for thorough statistical analysis.

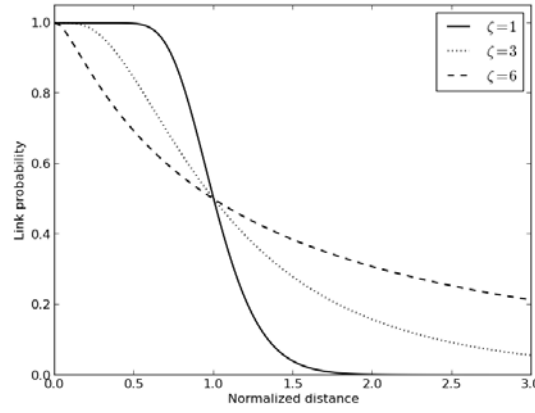


Figure 5.10: Link probability versus the normalised distance.

2. Therefore the result reported in Table 5.2 agrees with our conjecture that if $\zeta \geq 6$ the border effect is negligible unless ρ is extremely large.

As we have seen in §.2.3, major reasons that the random graph model is not used to model wireless *ad hoc* networks is that links in random graphs are not correlated and do not depend on distance.

Links correlation can be reflected by the *clustering coefficient*. The clustering coefficient, C , of a graph, $G = (V, E)$, is defined as follows (Watts and Strogatz, 1998):

$$C = \frac{1}{|V|} \sum_{v \in V} C_v, \quad (5.3.1)$$

where C_v is the *local clustering coefficient* of node v and is defined as follows: suppose that node v has k_v neighbours. Then there are $k_v(k_v - 1)/2$ possible links among the neighbours of v . We define C_v as the ratio between the actual *existent* links to all possible links. Let's define C_v formally, let N_v be the set of neighbours of v . Then

$$C_v = \frac{2|E \cap N_v \times N_v|}{k_v(k_v - 1)}, \quad (5.3.2)$$

Figure 5.11 helps in understanding the local clustering coefficient.

The clustering coefficient can be interpreted as the probability that two randomly chosen nodes are connected given that they have a common neighbour. By now, it should be clear that the clustering coefficient reflects the correlation between links. A high clustering coefficient corresponds to a highly link-correlated graph. In the context of social networks, the clustering coefficient measures the cliquishness of typical friendship circles and the local clustering coefficient answers the question: what is the probability that two randomly chosen friends of v know each other?

Anyway, we regard the clustering coefficient as a measure of correlation between links. A random geometric graph has a high clustering coefficient

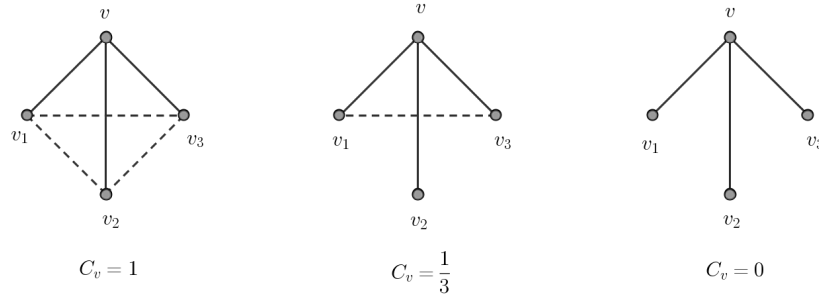


Figure 5.11: Local clustering coefficient of node v .

while a random graph has a low one. The clustering coefficient of wireless *ad hoc* networks, depending on ζ , is somewhere in the middle. To investigate the effect of ζ on the clustering coefficient of log-normal graphs, we calculate the clustering coefficients for the three graphs of interest and plot them against the average degree as shown in Figure 5.12. From the figure we notice that the curve representing the clustering coefficients of log-normal graphs approaches that of random graphs as ζ increases. When $\zeta = 6$ the two curves almost merge. This implies when $\zeta \geq 6$ links in log-normal graphs are weakly correlated and thus the log-normal model is not a good model for wireless *ad hoc* networks in this case.

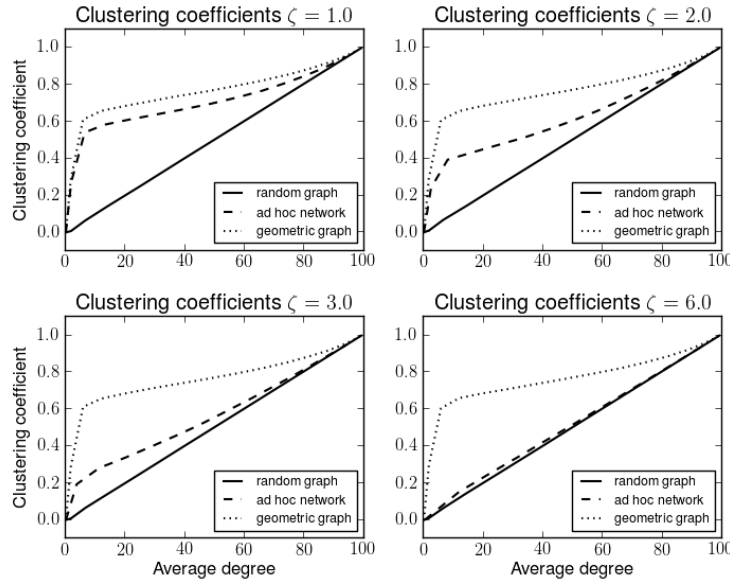


Figure 5.12: Clustering coefficients for random geometric graphs, log-normal graphs and random graphs plotted against the average degree.

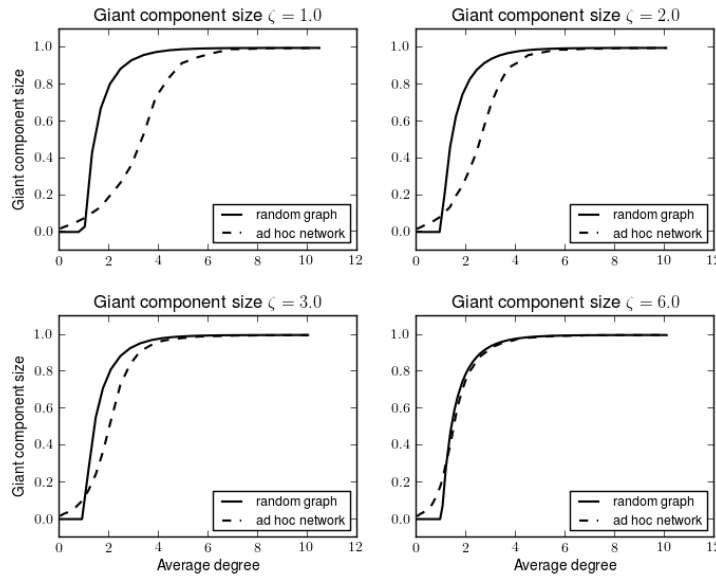


Figure 5.13: Giant component size as a fraction of the total network size for random graphs and log-normal graphs plotted against the average degree.

The last respect that we use to study the effect of increasing ζ on the reliability of the log-normal graph in modelling wireless *ad hoc* networks is the *giant component size*. Giant components are one of the most important features of random graphs and have been studied very well. For random graphs, if n is large, it has been shown (Hekmat and Van Mieghem, 2006) that the size of the giant component is the non-zero solution of the following equation:

$$C = 1 - \exp(-\bar{D} C), \quad (5.3.3)$$

where \bar{D} is the average degree. We compare the size of the giant component of random graphs obtained by numerically solving Eq.(5.3.3) with that of log-normal graphs obtained via simulation for different values of ζ . Figure 5.13 shows the result when $|A| = 100$ and $n = 100$.⁹ From the figure we can see that as ζ increases the difference between the giant component sizes of random graphs and log-normal graphs decreases. When $\zeta = 6$ they are almost the same.

Based on the above argument we conclude that the log-normal model is not a realistic representation of wireless *ad hoc* networks if $\zeta \geq 6$.

⁹We do the experiment for different values of n but not reported here. All results agree with that shown in Figure 5.13.

Chapter 6

Conclusion

6.1 Summary

We have seen that modelling signal propagation is essential for modelling wireless *ad hoc* networks reliably. Unfortunately modelling signal propagation is a very difficult task due to the various mechanisms that govern the propagation of signals. Through the path loss model and the log-normal model the strength of the received signal can be well approximated. The former model ignores the variations in the received signal strength whereas the latter model takes them into consideration. A wireless *ad hoc* network is equivalent to a random geometric graph if the path loss model is used and it is a completely new graph which we refer to as the log-normal graph if the log-normal model is used. In random geometric graphs, the occurrence of the link event is deterministic while it is probabilistic in log-normal graphs.

A random geometric graph and a log-normal graph model the situation in which there is no topology control. Consequently energy consumption and interference are very high. We have presented the XTC and CBTC algorithms which run on a random geometric graph or a log-normal graph to compute a subgraph which has certain desired properties. The algorithms clearly confine node energy consumption by restricting node transmitting ranges and we have shown through simulation that they significantly reduce interference.

The XTC is the cheapest in terms of the information it requires to operate correctly. It is based on neighbourhood information. If it runs on a random geometric graph it computes a variant of the relative neighbourhood graph locally in a distributed manner. It maintains connectivity. Although it is not a length spanner, we have shown through simulation that practically its length stretch factor does not grow as the number of nodes.

The CBTC relies on direction information to build a variant of the Yao graph in a fully distributed and local manner. It is proved it preserves connectivity if $\theta \leq 5\pi/6$ where θ is the algorithm's parameter. Moreover if $\theta \leq 2\pi/3$, asymmetric links can be removed without violating connectivity. We have

proved that it is a power spanner if $\theta \leq \pi/2$.

We have compared XTC and CBTC in terms of interference reduction, length stretch factor and maximum degree. We have not been able to preferred one over the other because both perform well in terms of interference reduction, their length stretch factor is not high and they maintain a constant maximum degree. However when the asymmetric links are not removed, XTC clearly outperforms the CBTC in terms of interference reduction.

We have chosen to study the XTC and CBTC algorithms because they are very practical and maintain connectivity. Their competitive algorithms either only ensure connectivity w.h.p or not practical by requiring expensive information.

Finally, we would like to identify the novelty of this work. From the analytical perspective, we have proved that CBTC is a power spanner and XTC maintains connectivity. In addition, we have demonstrated a contradiction in the proof of a power stretch factor obtained for the Yao graph in (Li *et al.*, 2001). Furthermore, we have compared this power stretch factor with correctly obtained power stretch factors to show that it behaves differently and not as expected. From the simulation perspective, we have obtained three important results. Firstly, we have shown that the degree distribution of wireless *ad hoc* networks modelled as log-normal graphs is always binomial in the absence of border effect and it is binomial in the presence of the border effect as long as the average degree is not high. We find that there is no fixed threshold for the average degree at which the degree distribution is no longer binomial. It depends on the network characteristics and the surrounding environment. Secondly, we have shown that connectivity can be guaranteed w.h.p if the minimum node degree is 1. Last but not least, we have shown that the log-normal model is not a realistic representation of wireless *ad hoc* networks if the environmental parameter is at least 6.

This work could have been improved in various ways which we shall briefly mention in the subsequent section.

6.2 Future work

Although we have shown through simulation that XTC and CBTC significantly reduce interference, they are not guaranteed to produce a non trivial approximation to the optimal-interference topology. Sadly interference is ignored in almost all topology control protocols because of the common conjecture that interference reduction is a direct consequence of controlling node transmitting ranges. The only exception is (Burkhart *et al.*, 2004) which proposes three protocols that produce an optimal-interference topology.

Many topology control algorithms with different novel aspects have been proposed in the last decade. Therefore a comparative study is needed to fill in this gap. In particular, we are interested in evaluating XTC and CBTC,

through a network simulator, in terms of the performance of *ad hoc on-demand distance vector routing*, network lifetime, network throughput, robustness to dynamism, etc.

A typical topology control algorithm restricts node transmitting ranges by somehow ignoring far away neighbours in favour of closer neighbours. This is a greedy view because each node tries to save its own power by shortening its transmitting range. We suggest an alternative approach which combines the greedy view and a network-wide view. To introduce the approach let us describe a node with battery power less than a predetermined threshold as *consumed* and otherwise as *active*. The alternative approach *somehow* connects each node to a nearest neighbour as long as it is active. If it is not, it considers the next nearest neighbour and if it is active it connects to it; otherwise it keeps searching for a near active neighbour. Obviously, this approach might lead to emergence of very long links whose costs are very high. To avoid this problem we can define another threshold on the cost of wireless links. If the cost threshold is exceeded, the node will connect to its most nearest neighbour regardless of its battery power. In this case a consumed neighbour is preferred to an active neighbour whose link cost exceeds the cost threshold. For the consumed nodes we strengthen the condition of neighbour selection to make them have less neighbours and consequently less overhead. Let us illustrate this approach by using a variation of CBTC as an example. Consider a generic active node. In each cone of angle θ around the active node we connect the node to a nearest neighbour as long as it is active. If it is not, the active node searches in the same cone for a next nearest neighbour. If it is active it connects to it and otherwise it keeps searching. If there is no active neighbour whose link cost does not exceed the cost threshold, it connects to the most nearest consumed neighbour in that cone. Now consider a generic consumed node. It does the same as the generic active node except in this case we strengthen the criterion of choosing a neighbour so instead of creating cones of angle θ , the consumed node will create cones of angle α where $\alpha > \theta$. As a result it will have less neighbours than active nodes on average. This approach will definitely complicate the analysis of connectivity, yet it is a new perspective that we think is worth trying.

In homogeneous topology control we have only done simulation. The analytical treatment is delayed for future investigation. Before (Bettstetter and Hartmann, 2005) and (Ta *et al.*, 2009) all the analyses of connectivity is based on the random geometric graph model which is certainly a very simple abstraction of reality. Although very rigours, the work done in (Bettstetter and Hartmann, 2005) and (Ta *et al.*, 2009) can still be improved and made more practical by considering the border effect in both and considering a small number of nodes in (Ta *et al.*, 2009) instead of the asymptotic analysis of connectivity presented therein.

Bibliography

- Andersen, J.B., Rappaport, T.S. and Yoshida, S. (1995). Propagation measurements and models for wireless communications channels. *Communications Magazine, IEEE*, vol. 33, no. 1, pp. 42–49.
- Bettstetter, C. (2002). On the minimum node degree and connectivity of a wireless multihop network. In: *Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking & computing*, pp. 80–91. ACM.
- Bettstetter, C. (2004a). Failure-resilient ad hoc and sensor networks in a shadow fading environment. In: *IEEE/IFIP International Conference on Dependable Systems and Networks*.
- Bettstetter, C. (2004b). On the connectivity of ad hoc networks. *The computer journal*, vol. 47, no. 4, pp. 432–447.
- Bettstetter, C. and Hartmann, C. (2005). Connectivity of wireless multihop networks in a shadow fading environment. *Wireless Networks*, vol. 11, no. 5, pp. 571–579.
- Blough, D.M., Leoncini, M., Resta, G. and Santi, P. (2003). The k-neigh protocol for symmetric topology control in ad hoc networks. In: *Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, pp. 141–152. ACM.
- Bollobas, B. (2001). *Random Graphs (Cambridge Studies in Advanced Mathematics)*. Cambridge University Press.
- Bose, P., Maheshwari, A., Narasimhan, G., Smid, M. and Zeh, N. (2004). Approximating geometric bottleneck shortest paths. *Computational Geometry*, vol. 29, no. 3, pp. 233–249.
- Buchin, K. and Buchin, M. (2007). Topology control. In: Wagner, D. and Wattenhofer, R. (eds.), *Algorithms for Sensor and Ad Hoc Networks*, vol. 4621 of *Lecture Notes in Computer Science*, pp. 81–98. Springer Berlin Heidelberg.
- Burkhart, M., von Rickenbach, P., Wattenhofer, R. and Zollinger, A. (2004). Does topology control reduce interference? In: *Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing*, pp. 9–19. ACM.
- Cormen, T.H., Leiserson, C.E., Rivest, R.L. and Stein, C. (2001). *Introduction to Algorithms, Second Edition*. The MIT Press.

- Crawshaw, J. and Chambers, J. (2001). *A Concise Course in Advanced Level Statistics: With Worked Examples*. Nelson Thornes.
- Diestel, R. (2005). *Graph Theory (Graduate Texts in Mathematics)*. Springer.
- Glauche, I., Krause, W., Sollacher, R. and Greiner, M. (2003). Continuum percolation of wireless ad hoc communication networks. *Physica A: Statistical Mechanics and its Applications*, vol. 325, no. 3, pp. 577–600.
- Gupta, P. and Kumar, P.R. (1998). Critical power for asymptotic connectivity in wireless networks. In: *Stochastic analysis, control, optimization and applications*, pp. 547–566. Springer.
- Hekmat, R. (2006a). *Ad Hoc Networks: Fundamental Properties And Network Topologies*. Springer.
- Hekmat, R. (2006b). *Ad-hoc Networks: Fundamental Properties and Network Topologies*. Springer.
- Hekmat, R. and Van Mieghem, P. (2006). Connectivity in wireless ad-hoc networks with a log-normal radio model. *Mobile Networks and Applications*, vol. 11, no. 3, pp. 351–360.
- Labrador, M. and Wightman, P. (2009). *Topology Control in Wireless Sensor Networks: with a companion simulation tool for teaching and research*. Springer.
- Li, L., Halpern, J.Y., Bahl, P., Wang, Y.-M. and Wattenhofer, R. (2005). A cone-based distributed topology-control algorithm for wireless multi-hop networks. *Networking, IEEE/ACM Transactions on*, vol. 13, no. 1, pp. 147–159.
- Li, N., Hou, J.C. and Sha, L. (2003). Design and analysis of an mst-based topology control algorithm. In: *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 3, pp. 1702–1712. IEEE.
- Li, X.-Y., Wan, P.-J. and Wang, Y. (2001). Power efficient and sparse spanner for wireless ad hoc networks. In: *Computer Communications and Networks, 2001. Proceedings. Tenth International Conference on*, pp. 564–567. IEEE.
- Moaveni-Nejad, K. and Li, X.-Y. (2005). Low-interference topology control for wireless ad hoc networks. *Ad Hoc & Sensor Wireless Networks*, vol. 1, no. 1-2, pp. 41–64.
- National Institute of Standards & Technology, U.D.o.C. (2008 November). Nist/sematech e-handbook of statistical methods, <http://www.itl.nist.gov/div898/handbook/>.
- O’ Rourke, J. (2010). The yao graph y_6 is a spanner. *CoRR abs/1003.3713*.
- Oliveira, C.A. and Pardalos, P.M. (2011). *Mathematical Aspects of Network Routing Optimization (Springer Optimization and Its Applications)*. Springer.

- Ott, R.L. and Longnecker, M.T. (2008). *An Introduction to Statistical Methods and Data Analysis*. Cengage Learning.
- Penrose, M.D. (1999). On k-connectivity for a geometric random graph. *Random Structures & Algorithms*, vol. 15, no. 2, pp. 145–164.
- Rajaraman, R. (2002). Topology control and routing in ad hoc networks: a survey. *ACM SIGACT News*, vol. 33, no. 2, pp. 60–73.
- Rodoplu, V. and Meng, T.H. (1999). Minimum energy mobile wireless networks. *Selected Areas in Communications, IEEE Journal on*, vol. 17, no. 8, pp. 1333–1344.
- Santi, P. (2005a). The critical transmitting range for connectivity in mobile ad hoc networks. *Mobile Computing, IEEE Transactions on*, vol. 4, no. 3, pp. 310–317.
- Santi, P. (2005b). Topology control in wireless ad hoc and sensor networks. *ACM Computing Surveys (CSUR)*, vol. 37, no. 2, pp. 164–194.
- Santi, P. (2005c). *Topology Control in Wireless Ad Hoc and Sensor Networks*. Wiley. ISBN 9780470094549.
- Santi, P. and Blough, D.M. (2003). The critical transmitting range for connectivity in sparse wireless ad hoc networks. *Mobile Computing, IEEE Transactions on*, vol. 2, no. 1, pp. 25–39.
- Song, W.-Z., Wang, Y., Li, X.-Y. and Frieder, O. (2005). Localized algorithms for energy efficient topology in wireless ad hoc networks. *Mobile Networks and Applications*, vol. 10, no. 6, pp. 911–923.
- Spiegel, M.R., Schiller, J.J. and Srinivasan, R.A. (2000). *Schaum's Outline: Probability and Statistics, Second Edition*. McGraw-Hill.
- Stuedi, P., Chinellato, O. and Alonso, G. (2005). Connectivity in the presence of shadowing in 802.11 ad hoc networks. In: *Wireless Communications and Networking Conference, 2005 IEEE*, vol. 4, pp. 2225–2230. IEEE.
- Ta, X., Mao, G. and Anderson, B.D. (2009). On the giant component of wireless multihop networks in the presence of shadowing. *Vehicular Technology, IEEE Transactions on*, vol. 58, no. 9, pp. 5152–5163.
- Volbert, K. (2005). *Geometric Spanners for Topology Control in Wireless Networks*. Ph.D. thesis, Faculty of electrical engineering, computer science and mathematics, Institute of computer science and Heinz Nixdorf Institute, University of Paderborn.
- Wan, P.-J. and Yi, C.-W. (2004). Asymptotic critical transmission radius and critical neighbor number for k-connectivity in wireless ad hoc networks. In: *Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing*, pp. 1–8. ACM.

- Wattenhofer, R., Li, L., Bahl, P. and Wang, Y.-M. (2001). Distributed topology control for power efficient operation in multihop wireless ad hoc networks. In: *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3, pp. 1388–1397. IEEE.
- Wattenhofer, R. and Zollinger, A. (2004). Xtc: A practical topology control algorithm for ad-hoc networks. In: *Parallel and Distributed Processing Symposium, 2004. Proceedings. 18th International*, p. 216. IEEE.
- Watts, D.J. and Strogatz, S.H. (1998). Collective dynamics of small-world networks. *nature*, vol. 393, no. 6684, pp. 440–442.
- Wiki01, 2013 (Accessed on September 2013). Random geometric graph. Wikipedia, the Free Encyclopedia.
Available at: http://en.wikipedia.org/wiki/Random_geometric_graph
- Wiki02, 2013 (Accessed on September 2013). Kolmogorov-smirnov test. Wikipedia, the Free Encyclopedia.
Available at: http://en.wikipedia.org/wiki/Kolmogorov-Smirnov_test
- Wiki03, 2013 (Accessed on September 2013). Dijkstra's algorithm. Wikipedia, the Free Encyclopedia.
Available at: http://en.wikipedia.org/wiki/Dijkstra's_algorithm
- Wiki04, 2013 (Accessed on September 2013). Topology control. Wikipedia, the Free Encyclopedia.
Available at: http://en.wikipedia.org/wiki/Topology_control
- Woo, A., Tong, T. and Culler, D. (2003). Taming the underlying challenges of reliable multihop routing in sensor networks. In: *Proceedings of the 1st international conference on Embedded networked sensor systems*, pp. 14–27. ACM.
- Xiang-Yang, L., Song, W.-Z. and Wang, Y. (2005). Efficient topology control for ad-hoc wireless networks with non-uniform transmission ranges. *Wireless Networks*, vol. 11, no. 3, pp. 255–264.
- Xue, F. and Kumar, P.R. (2004). The number of neighbors needed for connectivity of wireless networks. *Wireless networks*, vol. 10, no. 2, pp. 169–181.
- Yao, A.C.-C. (1982). On constructing minimum spanning trees in k-dimensional spaces and related problems. *SIAM Journal on Computing*, vol. 11, no. 4, pp. 721–736.